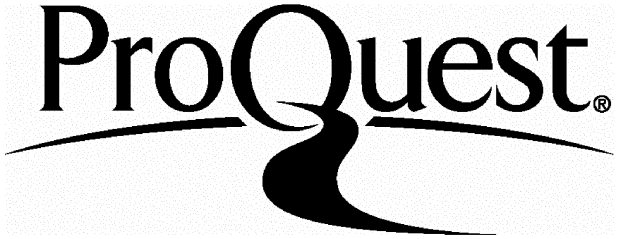
THE PERCEPTION OF SECOND-ORDER MOTION.

by Christopher Philip Benton

Submitted for the degree of Doctor of Philosophy

University College London



Declaration.

The work submitted here is the result
of the candidates own investigations.

Acknowledgements.

This work would not have been completed were it not for the excellent guidance provided by Alan Johnston. Additional thanks go to Peter McOwan (who unknowingly provided the role of second supervisor) and to the other various inhabitants of the vision lab, Willie Curran, Colin Clifford and Justin O'Brien, who all, in some way, left their mark on this thesis. Finally, to Diana for her support and forbearance.

Abstract.

In this thesis the notion of an independent non-linear channel for the perception of second-order motion is investigated. An examination of speed discrimination thresholds for first- and second-order bars and edges showed no differences in the patterns of response over changes in the temporal and spatial parameters of the stimuli. The higher thresholds for second-order stimuli may be accounted for by appealing to the properties of their noise carriers.

In a study of the direction of motion in reversed-phi stimuli, it was shown that luminance and contrast defined stimuli could elicit both forwards and reversed motion. The forwards motion in the contrast defined stimulus cannot be explained by the operation of a first- or second-order channel. Perception of motion direction in the contrast defined stimulus was dependent upon the characteristics of the noise carrier. Similar dependencies were observed when noise was added to the first-order stimulus. When the effect of carrier is taken into account, both types of stimulus show similar patterns of response over spatial and temporal frequency.

Modulation depth tuning curves for the detection of motion direction in stimuli where motion was defined by contrast and luminance microstructure were also investigated. Luminance microstructure can affect perceived contrast in contrast defined motion and also in static noise patterns. This implies that some early non-linearity exists in the human visual system. The filtering and rectification approach to recovery of the second-order motion should be highly effective at recovering the modulant. However, an estimate of the size of the non-linearity shows that only a relatively small distortion was necessary to account for the modulation depth tuning curves.

The results from this thesis indicate that the carrier is crucial to the perception of second-order motion. Differences in response to first- and second-order motion may depend upon properties of the stimuli rather than the operation of separate mechanisms. It is argued that the results cast some doubt over the notion of a second-order channel. A number of alternatives are discussed.

Table of Contents.

Chapter One

Introduction.

1. Introduction.

We depend on motion information for a wide variety of tasks, from simply helping us to keep our balance (Lee & Lishman, 1975) through to the high level extraction of three dimensional structure of an object or scene from its component velocities (Braunstein, 1994). The recognition of motion is an important part of the behaviour of a huge number of creatures from scout bees communicating the position of a food source via their "waggling dance" (Frisch, 1967) to humans understanding of sign language, lip reading and gestures. When all that is visible of a walking person is a number of reflective bands attached to their joints and extremities, we can instantly recognise that moving pattern as a walking person, even though the static pattern may be meaningless (Johannson, 1973). Motion itself signals some element of change, the very fact of its existence can potentially tell us something new about our visual world. The images falling across our retinae are inherently spatio-temporally structured. To extract information from this shifting optic array our visual system must be sensitive not only to spatial patterns but also to temporal change.

Given the importance of motion as a source of information, it would seem reasonable to assume that some elements of the visual system are specifically designed to measure motion. Evidence for this is provided by the motion aftereffect, an early example of which is the waterfall illusion described by Wohlgemuth (1911) (also see Verstraten, 1996). If one fixates on some translating texture (e.g. a sine wave grating), such that the texture translates across the point of fixation, then, when the texture stops moving, a strong percept of motion is seen in the reverse direction. The perceptual dissociation of position and motion implies that the motion is not being extracted from simple positional cues. Motion and position must be regarded as separate sensory dimensions (Gregory, 1966). The existence of the motion aftereffect implies very strongly that the visual system contains mechanisms that are designed to be sensitive to motion itself. Furthermore, the study of the anatomy and physiology of the visual of animals and humans has shown that there exist mechanisms dedicated to the extraction of motion.

1.1. Overview of Anatomy and Physiology.

The purpose of this section is twofold. Firstly, to provide a simple overview within which to place some of the discussion that occurs later in this chapter. Secondly, to provide an overview of evidence suggesting the existence of a functionally separate and anatomically distinct region of the brain dedicated to motion processing. Good reviews of the subject are provided by Celesia & DeMarco (1994) and Zeki (1993).

There are two types of photoreceptor present in the retina: rods and cones. Rods operate in low lighting conditions (ie. night vision) whilst cones operate at scotopic levels of luminance (ie. normal vision). Cones can be split into a further 3 types depending upon which type of photopigment they contain. Because of the different photopigments these three types of cone are preferentially activated by different wavelengths of light. They can be described as short-wave sensitive, middle-wave sensitive and long-wave sensitive and show peak responses to wavelengths of 430nm, 530nm and 555nm respectively. By comparing the outputs of cones tuned to different wavelength, the visual system can estimate the wavelength of the light falling on the cones.

Information from the photoreceptors is conveyed to ganglion cells through a complex set of interconnections between a number of other retinal cells (bipolar cells, horizontal cells and amacrine cells). In the primate visual system the majority of ganglion cells display a centre surround organization and may be described as on-centre or off-centre. An on-centre cell will have an excitatory centre and an inhibitory surround, and vice versa for an off-centre cell. The spatial cross-section of the response profiles of these cells have traditionally been modelled by two a dimensional difference of Gaussians (Rodieck, 1965; Soodak, 1986; Smith, Chino, Ridder, Kitagawa & Langston, 1990), but a two dimensional Laplacian of Gaussian function also provides a reasonable fit (Marr & Hildreth, 1980; Marr & Ullman, 1981). As the output of a cell is always positive, the use of on-centre and off-centre cells provides a way for the visual system to encode both local increments and local decrements in illumination. The output from the ganglion cells may be split into ON and OFF pathways on the basis of this description.

The majority of ganglion cells project to the lateral geniculate nucleus (LGN). The area of the LGN to which they project has provided another way of classifying the retinal output into two pathways - the M and P pathways. M-cells in the retina project to the magnocellular layers of the LGN, P-cells project to the parvocellular layers. The latter are colour sensitive and demonstrate either red-green or blue-yellow spectral opponency. M-cells respond to luminance regardless of wavelength. The generally held view is that little or no processing of the visual information occurs in the LGN and that this area simply transmits signals to the primary visual cortex, also termed area V1.

The M/P distinction is one of the cornerstones of current thinking about the way the primate visual system is organized. To a certain degree these pathways appear to continue to remain independent from their source in the retina up through the LGN, onwards through the striate cortex, and then into different locations in the extra-striate cortex. The M-pathway then projects into areas of the parietal lobe whilst the P-pathway projects into areas of the temporal lobe. Cells in the magnocellular layers of the LGN project to layer 4B of the primary visual cortex. Cells in this region have projections up to the middle temporal area (also termed V5) both directly and indirectly via an area of secondary visual cortex, V2. Whilst areas V1 and V2 appear functionally heterogenous, a great deal of evidence suggests that cells in extra-striate area V5 respond to motion and the vast majority appear to be directionally selective (Zeki, 1974; van Essen, Maunsell & Bixby, 1981; Albright, 1984; Newsome, Wurtz, Dürsteler & Mikami, 1985; Newsome & Paré, 1988).

The findings in non-human primates (mainly macaque) are generally held to apply to human visual processing. Evidence of functional specialisation for motion processing in humans is provided by Zihl, von Cramon & Mai (1983) who studied a patient with lesions to the extra-striate cortex. The patient demonstrated profound akinotopsia, an inability to perceive motion, but other aspects of her visual perception were apparently normal. A homologue of macaque V5 has been identified in the human brain by the use of positron emission tomography (PET) (Zeki, Watson, Lueck, Friston, Kennard & Frackowiak, 1991; Watson, Myers, Frackowiak, Hajnal, Woods, Mazziotta, Shipp & Zeki, 1993). Watson *et al* coregistered their PET images with images of the brains of their subjects obtained using magnetic resonance imaging (MRI). The greater resolution of MRI allowed Watson *et al* to identify human V5 as an area previously identified by Flechsig and termed *Feld 16*. The

existence of human V5 has also been confirmed by MRI (Tootell, Reppas, Kwong, Malach, Born, Brady, Rosen & Belliveau, 1995) and magneto-encephalography (SQUID) (Anderson, Holliday, Singh & Harding, 1996).

In an novel paradigm, Beckers & Zeki (1995) looked at the effects of inactivating human V5 using transcranial magnetic stimulation (TMS) on awake behaving volunteers (see also Hotson, Braun, Herzberg & Boman, 1994). They found that stimulation of V5 abolished the perception of motion. It is interesting to note that TMS applied to V1 did not totally abolish the perception of motion, Beckers & Zeki claim that this may be due to a fast neural pathway that projects from the retina to V5 without passing through V1. It should also be noted that TMS is a relatively new and untried technique. It is difficult to say to what degree and over what precise region of the brain the disruption occurs. However their induced akinotopsia caused by TMS to the extra-striate cortex fits well with the Zihl *et al*'s (1983) study of the akinotoptic patient.

In summary, there appear to be two parallel pathways, the P pathway and the M pathway. That latter projects to area MT/V5. A great deal of evidence strongly suggests that this area plays a crucial part in motion perception. However, there are a great deal of interconnections between and within areas of the visual pathway that can potentially exchange information between the M and P pathways. Whilst the notion of two parallel pathways does appear to capture something fundamental about the visual system, there are a number of pieces of evidence that do not totally concur with this view (for a review see Merigan & Maunsell, 1993). For example, patient L.M., originally studied by Zihl *et al* (1983) has been shown to perceive global coherent motion in random dot motion stimuli, though her performance is highly susceptible to noise (Baker, Hess & Zihl, 1991). She can also extract 3D structure from motion (Rizzo, Nawrot & Zihl, 1995). Merigan, Byrne & Maunsell (1991) found that motion perception in primates was only partially affected by magnocellular lesions. In human psychophysics, an examination of perceived motion direction in peripherally presented stimuli concluded that a reversed motion effect was produced by undersampling of the P cell mosaic rather than the M cell mosaic (Anderson, Drasdo & Thompson, 1995). Other evidence comes from studies that show that cells in MT/V5 response to motion in isoluminant stimuli, however this latter finding is the subject

of some dispute (Gegenfurtner, Kiper, Beusmans, Carandini, Zaidi & Movshon, 1994; Dobkins & Albright, 1995).

1.2. Models of Motion processing.

Low level models of motion processing may defined as those models which seek to extract motion directly from spatiotemporal filtering operations applied to the luminance surface, rather than by extracting motion from some higher order representation, as in, for example, a feature tracking mechanism (see section 1.3.6 below). Whilst a huge number of motion models are included in the computational literature, those that have been specifically identified as potential models of biological motion processing can be divided into three broad approaches: correlation models, energy models and gradient models.

1.2.1. Correlation/Reichardt models.

The correlation model represents the earliest and simplest class of motion detection model applied to animal vision. The prototypical correlation model consists of two spatially separated receptors. The output from both of these is multiplicatively combined, but with a temporal delay on the output of one of the receptors preceding the multiplicative stage (Figure 1.1a). This arrangement will also respond to static patterns, for example if both of the receptors are continually excited. This problem can be avoided by the use of an opponent detection arrangement such as that described by Reichardt (1961) to account for optomotor responses in insects. A simplified version of this model is outlined in Figure 1.1b. As the majority of early work on correlation models was carried out by Reichardt and his co-workers, models of this type are often referred to as Reichardt models.

Clearly, there is some difficulty applying a model developed for insects to human beings. In primates, there appears little or no evidence that directional selectivity in cells occurs before the primary visual cortex. The model cannot therefore receive direct photoreceptor input. This in itself does not present any great problem for the correlation approach, it may simply be the case that the input to the model may be the outputs from the LGN or from other cells in V1. The basic correlation model does however have a problem with temporal and spatial aliasing (van Santen & Sperling, 1984). Whether or not aliasing may occur is dependent on

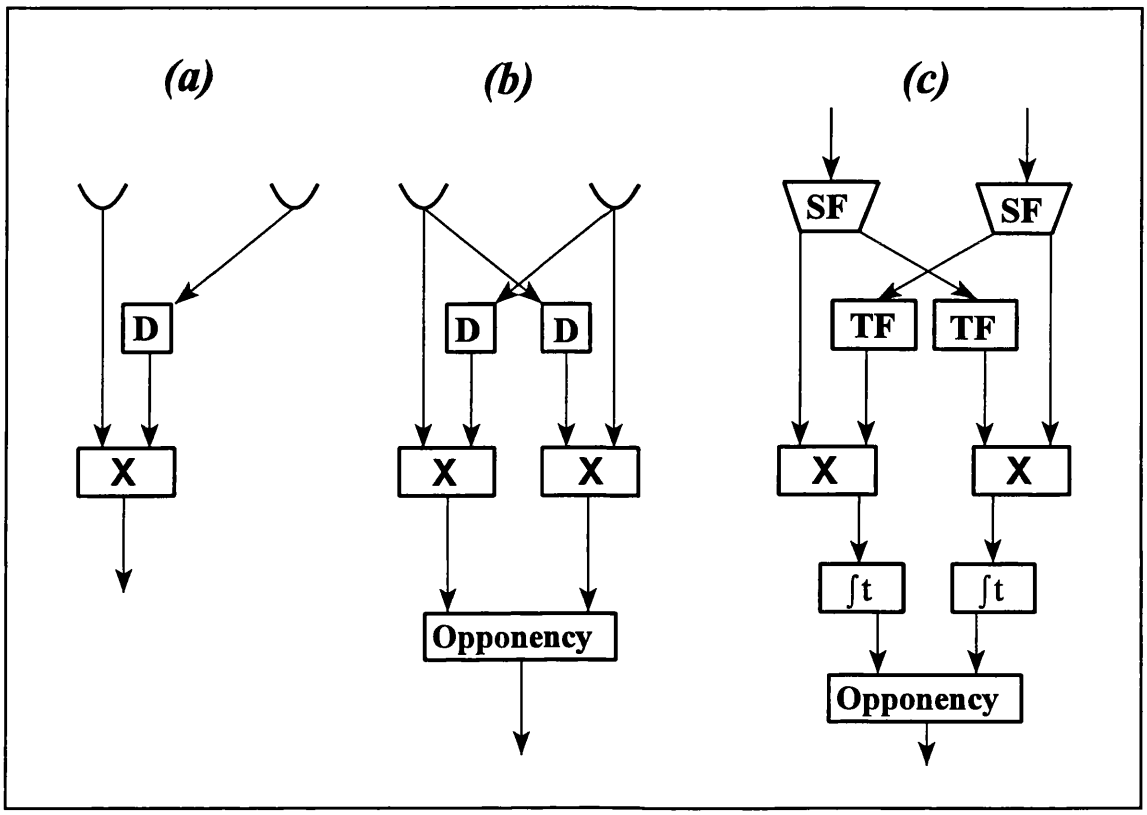


Figure 1.1: Correlation/Reichardt models. (a) basic subunit, **D** indicates temporal delay, **X** indicates multiplication.(b) subunits organized to give motion opponency. This latter stage is implemented by taking the difference between the subunits. (c) Elaborated Reichardt Model. **SF** indicates spatial filtering, **TF** indicates temporal filtering, $\int$**t** indicates temporal integration.

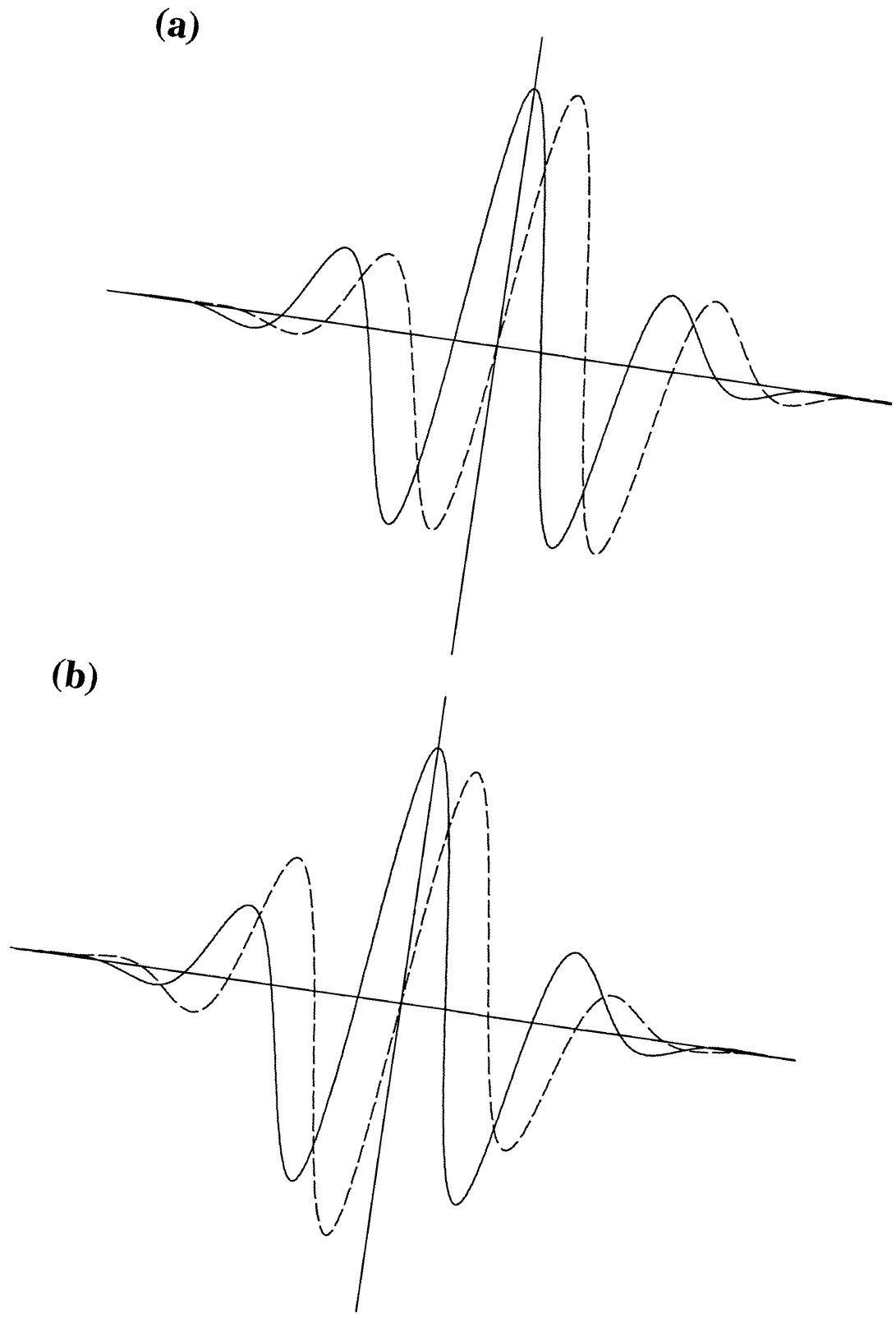


Figure 1.2: (a) Even Gabor functions separated by a quarter of a cycle, (b) Even and odd Gabor functions.

the relationship between the spatial and temporal frequencies in the input stimulus and the spatial separation and temporal delay/filtering within the model. One way that temporal aliasing may be avoided is to prevent the phase delay for any temporal frequency from being greater that half a temporal cycle. In their original extension of the Reichardt model, van Santen & Sperling propose a low pass temporal filter with a weighting function of the form $e^{-t/\tau}$, they show that a filter with this profile will guarantee that the phase delay is always between zero and $\pi/2$.

In order to avoid spatial aliasing van Santen & Sperling (1984) add a spatial filtering stage to the input of the model (Figure 1.1c). At this stage, their model is no different from that described by Reichardt (1961), however they show that by choosing the separation between the spatial filters such that they are separated by a phase difference of $\pi/2$, the problem of spatial aliasing can be avoided. This may be accomplished by a simple spatial separation (see Figure 1.2a) or by the use of odd and even Gabor filters (Figure 1.2b), a suggestion originally put forward by Watson & Ahumada (1983). In a later change to their model, which they term the "Elaborated Reichardt Detector" (ERD) van Santen & Sperling (1985) use Watson & Ahumada's (1983) suggestion of a temporal filter that approximates to a hyperbolic filter over a wide range of temporal frequencies. Watson & Ahumada (1983) show that a hyperbolic filter has a constant phase lag of $\pi/2$. The central importance of van Santen & Sperling's approach is not the changes they made to the Reichardt model. Indeed they simply brought together a number of ideas and suggestions that were already extant in the literature. However, they do provide a rigid mathematical characterisation of the performance of the model, and their model provides a generalisation of a model developed for explaining insect vision into a model of human low level motion perception.

1.2.2. The Energy Model.

Watson & Ahumada (1983, 1985) suggested that human motion perception may best be understood by viewing the phenomenon within the frequency domain, a perspective that has its precedent in the field of spatial vision (Robson, 1966; Campbell & Robson, 1968; Maffei & Fiorentini, 1973). The purpose of Watson & Ahumada's (1983) paper was to provide a brief look at the utility of a Fourier decomposition approach to the understanding of human motion perception. They proposed a linear model which is similar to the ERD model

described above except that the multiplicative stage is replaced by simple summation. A diagram of their model is shown in Figure 1.3. The hyperbolic filters simply act to convert odd functions into even and even functions into odd (see Watson & Ahumada, 1985). The model can best be understood in terms of its response to a translating sinusoidal grating. Whereas in one direction the outputs of the two channels will be in phase such that the output of the summed response is twice that of the response in one channel, when the grating translates in the opposite direction then the signals in the two channels will be π radians out of phase so that there is no overall summed response.

The Watson & Ahumada motion detector can be regarded as a detector of Fourier motion energy that is localised in frequency space. By analysing motion signals in the Fourier domain and providing a model of how biological motion detectors might act as detectors of Fourier energy, Watson & Ahumada provided the basis for the energy model approach to motion perception. Watson & Ahumada note that their model has a number of desirable characteristics. The model is selective for motion direction but does not encode velocity. Human velocity discrimination at threshold is quite poor (Watson & Robson, 1981). This may be accounted for if it is presumed that velocity is extracted by examining outputs across a number of motion energy detectors. Furthermore, the model is selective also for spatial frequency and orientation, with the preferred direction of motion orthogonal to the spatial orientation of the sensor. These properties accord with the findings of studies that have examined the response of cells in the primary visual cortex of primates (DeValois, Albrecht & Thorell, 1982; DeValois, Yund & Hepler, 1982).

Currently, the most influential model of low level biological motion processing is that proposed by Adelson & Bergen (1985, 1986). It is clear that the output of the Watson & Ahumada model detailed above is sinusoidal in response to a sinusoidal input. At any given time the output is dependent upon the phase relationship between the stimulus and the initial filters. Adelson & Bergen present a model that is phase independent by combining the outputs of two detectors similar to those described by Watson & Ahumada (1983, 1985). If we consider a Gabor filter oriented in space and time then the output of that to a translating sine wave is a sine wave. If we take another Gabor filter that is in quadrature to the first, then the output of this filter is again a sinusoid. However there is a phase difference of $\pi/2$ radians between the outputs of the two filters. If we square the outputs of both filters

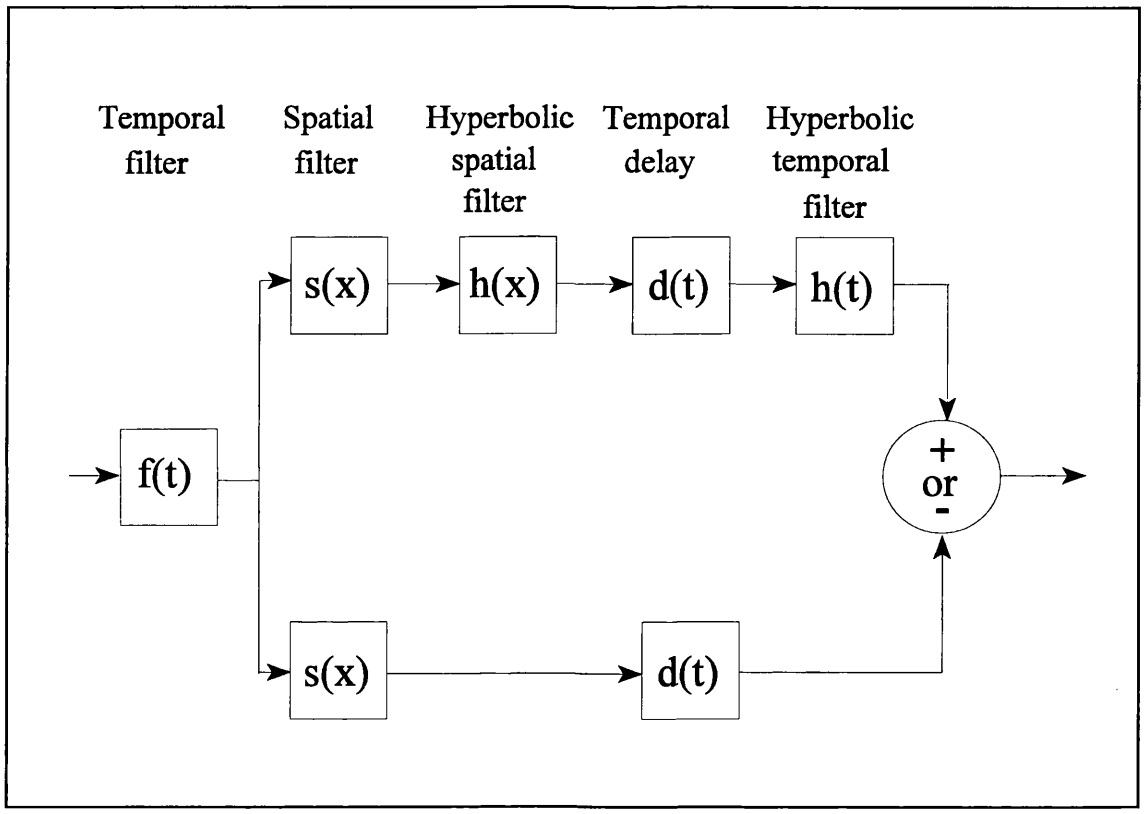


Figure 1.3: Motion detector described by Watson & Ahumada (1983).

and then add them, we obtain a phase invariant measure of Fourier motion energy. This procedure takes advantage of the fact that $\sin^2\theta + \cos^2\theta = 1$ (the phase difference between $\sin\theta$ and $\cos\theta$ is $\pi/2$ so this equation can be rewritten $\sin^2\theta + \sin^2(\theta + \pi/2) = 1$). Computationaly, spatio-temporally oriented Gabors may be implemented by the addition and subtraction of filters composed of space-time separable spatial and temporal Gabor functions. In this arrangements all spatial and temporal responses for the linear filters are $\pi/2$ out of phase with one another. The use of these filters therefore prevents both temporal and spatial aliasing from occurring.

Adelson & Bergen note that spatio-temporally oriented Gabors are probably not biologically plausible filters. They show how a set of four filters created from biologically plausible filter profiles can be combined to form an approximation to spatio-temporally oriented Gabors. In the spatial domain they utilise the first and second derivatives of a Gaussian, and in the temporal domain they use two impulse functions. Whilst the model does show a larger response for motion in one particular direction, it will also show a response for motion in the opposite direction and will also respond to a static pattern. In a similar fashion to the Reichardt model, Adelson & Bergen incorporate motion opponency into their model. The stages of the Adelson & Bergen model and space-time plots of their biologically plausible filters are shown in Figure 1.4. The notion of motion opponency sits well with the literature regarding the motion aftereffect (Wohlgemuth, 1911). It can be argued that, for example, a rightward signalling motion detector has been desensitised by the continual presentation of rightwards motion. When presented with a static pattern the leftward signalling detector will signal the larger response, leftward motion will be perceived.

The response of the energy model is not only affected by the spatial and temporal frequencies present in the input, but is also proportional to stimulus contrast. The van Santen & Sperling and Watson & Ahumada models are similarly affected by stimulus contrast. To get around this problem Adelson & Bergen suggest that a measure of static energy may be taken. By comparing the output of this channel with the outputs of the directional channels, Adelson & Bergen suggest that a contrast independent measure may be obtained. They make a direct analogy with the colour system whereby colour is represented by the output of three types of cone (see above). The scheme can be extended to cover a number of motion energy detectors within the same spatial frequency channel. As

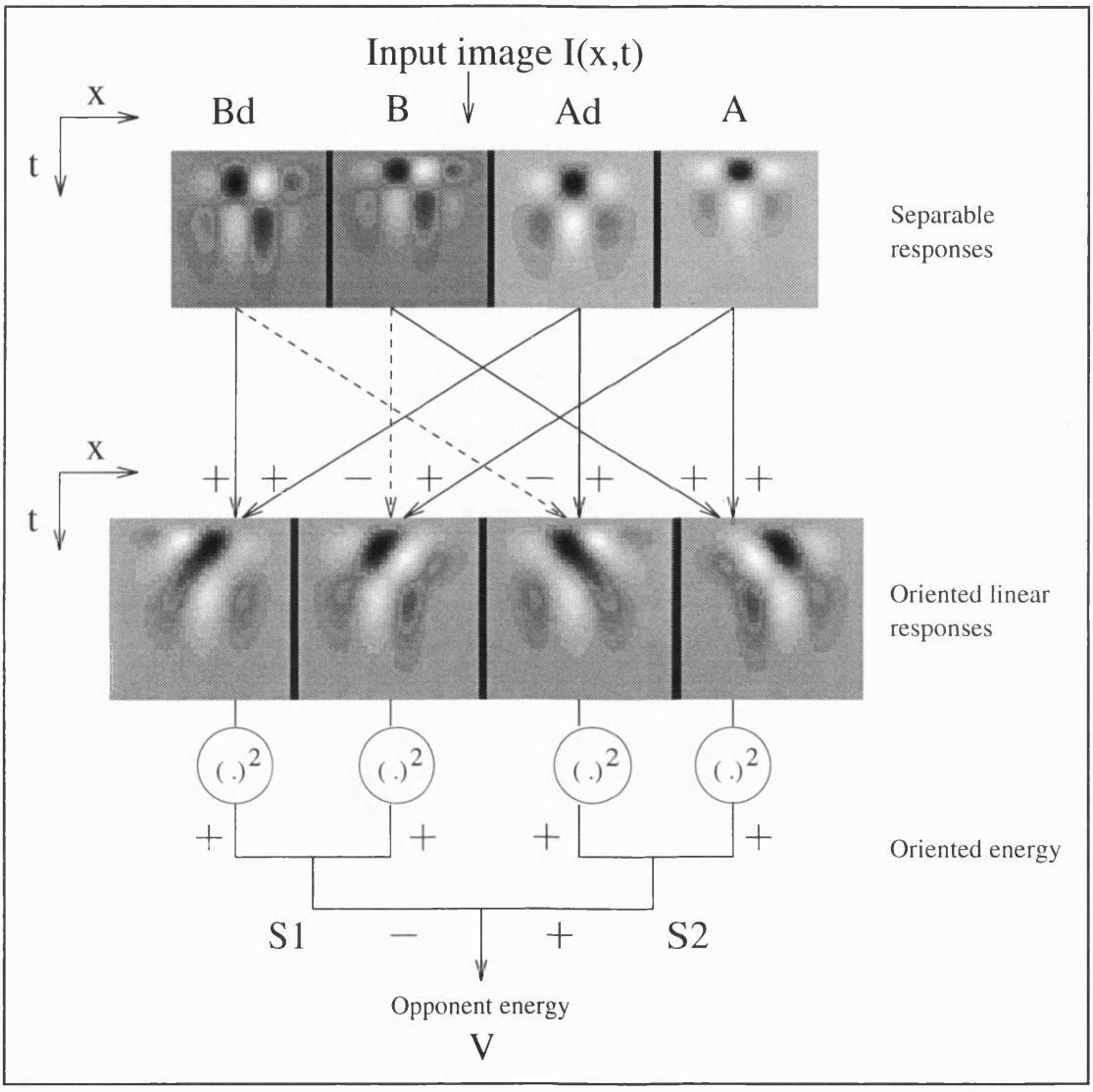


Figure 1.4: Schematic diagram showing energy model described by Adelson & Bergen (1985).

the comparison of outputs involves taking ratios, the measure of velocity may become increasingly unreliable as the denominator decreases. Adelson & Bergen suggest that measures of confidence may be attached to the velocity measures and that the outputs from a number of channels may be combined to recover a stable estimate of velocity. Additional schemes for the extraction of velocity from arrays of motion energy detectors have subsequently been advanced (ie. Heeger, 1987; Grzywacz & Yuille, 1990). It should be noted that in all cases the final estimate of velocity must be extracted from the combined outputs of a number of channels tuned to different spatial and temporal frequencies.

1.2.3. Gradient models.

Gradient models have one strong advantage over the other models described above; they provide a well defined account of how velocity may be recovered. Gradient methods are based on the following two observations. Firstly, when viewed in a space-time cube, the trajectory of a rigidly translating object forms oriented lines through that space. If we assume that there are no abrupt changes of lighting as the object moves through space and time, then we can assume that the luminance along these oriented lines is locally conserved. Secondly, the slope of the lines in the space-time cube is a measure of velocity. This is shown in Figure 1.5 Therefore if one can recover the gradient over which luminance is conserved, one can extract a local measure of velocity. The gradient can be recovered by taking the ratio of temporal change over the rate of spatial change. If we consider the process in only one spatial dimension and use for our example a space-time separable filter $g(x)h(t)$. From this filter we can derive another filter by taking the partial derivative with respect to the spatial dimension x to obtain $g'(x)h(t)$. We can derive another filter by taking the partial derivative with respect to the temporal dimension t, obtaining function $g(x)h'(t)$. The result of filtering the derivative of an image with a function is equivalent to filtering the image with the derivative of the function. If we take an input image $I(x,t)$ and filter it with the function $g'(x)h(t)$ we obtain the image $I_{1,0}(x,t)$. In this terminology $I_{n,p}(x,t)$ is the result of filtering the image $I(x,t)$ with the filter created by differentiating the function $g(x)h(t)$ n times with respect to x and p times with respect to t. If we take the input image $I(x,t)$ and filter it with the function $g(x)h'(t)$ we obtain the image $I_{0,1}(x,t)$. By dividing $I_{0,1}(x,t)$ by $I_{1,0}(x,t)$ we obtain local measures of image velocity (Fennema & Thompson, 1979; Johnston, McOwan & Buxton, 1992; McOwan & Johnston; 1996). In fact, in the terminology outlined

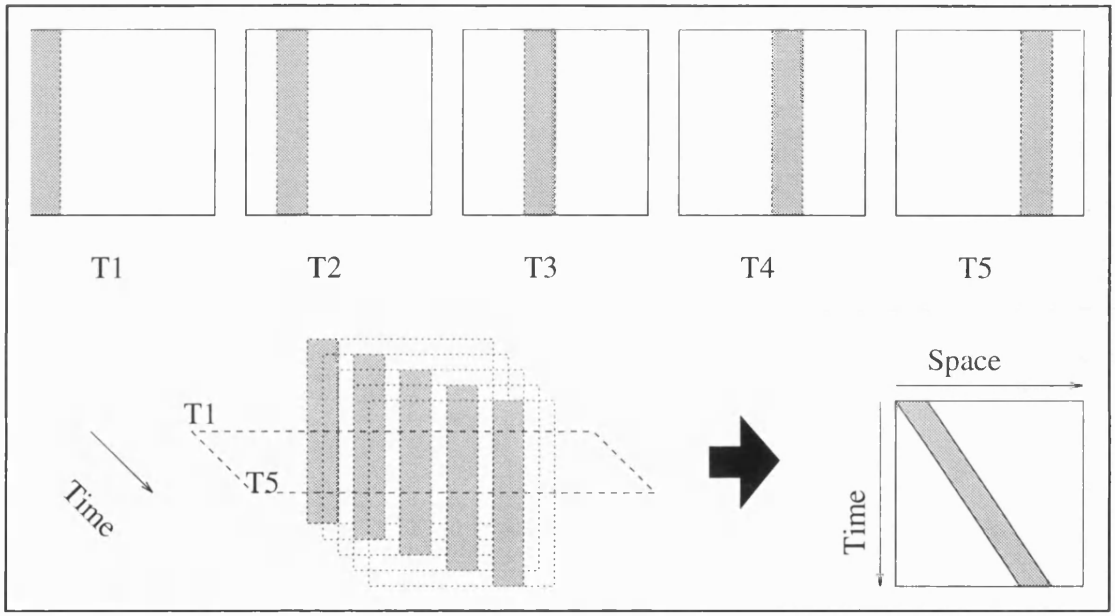


Figure 1.5: Motion as orientation in space and time. T1 to T5 represent frames taken from a motion sequence showing a translating grey bar. If we take a spatial slice through each frame and then stack the slices on top of one another, we obtain a grey line that is oriented in space and time.

above, velocity is described by the following equation (Fennema & Thompson, 1979; Horn & Schunk, 1980):

$$v = \frac{-I_{0,1}(x,t)}{I_{1,0}(x,t)}$$

Because gradient models involve the use of a quotient, they can produce an infinite response when the denominator is zero. This problem can be avoided by multiplying both the top and bottom of the equation by the denominator. This merely sets the output to zero when the denominator is zero. However, this is still an unsatisfactory state of affairs. The model may produce a zero response in an area where motion exists (ie. when the spatial derivative is zero). Another method by which the velocity calculation to be made more reliable is to combine a number of velocity measures by, for example, using a least means square procedure. This may be done by integrating over space (Lucas & Kanade, 1981) or by taking a number of other measures of velocity at each point. This latter approach has been implemented by Johnston *et al* (1992) who use increasing levels of spatial and temporal derivatives to stabilise the velocity measure.

1.2.4. Equivalence of motion models.

Van Santen & Sperling (1984, 1985) and Adelson & Bergen (1985,1986) examined the relationships between the various motion models. They prove that, for certain configurations, the energy model, the elaborated Reichardt model and the model developed by Watson & Ahumada (1983, 1985) are identical in terms of performance. They all behave as local detectors of Fourier motion energy. Adelson & Bergen (1986) showed that this could also be extended to a simple gradient model. Bruce, Green & Georgeson (1996) show that the energy model and the approach taken by Johnston *et al* (1992) may, in some cases, be mathematically equivalent. However it should be noted that this last derivation only applies to a limited subset of the Johnston *et al* model and it is not clear whether it can readily be generalised to the entire model.

In gradient models, to extract a measure of velocity, it is necessary to have a well defined mathematical relationships between the weighting functions of the spatio-temporal filters.

The gradient model extracts velocity precisely because it does have these relationships. At a very fundamental level it is these relationships that define the model as a gradient model. In derivations where the energy model is shown to be equivalent to gradient models, the similarity is achieved by utilising a set of filters that are ideal for the gradient model approach. These filters may not be ideal for an implementation of the energy model. In order to prove the equivalence between these two models the energy model essentially becomes a gradient model in that it succeeds in extracting velocity as a consequence of the mathematical relationships between the filter shapes. If energy models were truly identical to gradient models then there would be no need to address the issue of how to extract velocity from them (ie. Heeger, 1987; Grzywacz & Yuille, 1990).

1.2.5. Motion as oriented Fourier energy.

The idea that all low level motion models can be seen as detectors of Fourier motion energy has led to the tacit acceptance that models based upon linear filtering (ie. those described above) operate by detecting the orientation of lines or planes through the frequency domain origin that contain Fourier motion energy. This reason for this belief can be illustrated in the following manner. Velocity is given by the ratio of spatial frequency over temporal frequency. Therefore, in the frequency domain and considering only one spatial dimension, points that signify a particular velocity must lie on a line that passes through the origin.

To be more specific, velocity is given by the equation,

$$v = \frac{tf}{sf} .$$

If we choose a particular velocity v_1 and reformat the equation in the traditional $y = mx + c$ terms (where m is the slope and c is the intercept) then

$$tf = \left(\frac{1}{v_1} \right) . sf .$$

In other words, all the points in the frequency domain that describe a particular velocity must lie on a line through the origin with slope that is simply the inverse of velocity. If we

expand this up to two spatial dimensions, then points of equi-velocity all lie on a plane that passes through the frequency domain origin.

Although this is an extremely simple point and the description given above is rather laboured, this is because this concept is central to the way in which energy mechanisms (and by extension all other low level motion detection mechanisms) are seen to operate. If the contribution of stimulus contrast can be discounted, then at a fundamental level, energy models encode velocity simply as a function of their tuning characteristics. Of course this is not quite true because the response profile of an energy model in frequency space is not localised to a point but extends both spatially and temporally in frequency space. However, if we take a bank of motion energy detectors which overlap and fill an area of frequency space, then the velocity of a translating sine wave grating is given simply by the detector that gives the largest response (Fahle & Poggio, 1981).

The Adelson & Bergen (1985) model has proved so influential mainly because its operation can be explained in the simple heuristic terms outlined above. However, it should be noted that in reality the problem of extracting velocity from measures of motion energy is not facile. When the stimulus consists of a translating pattern with two or more components, then the response of any single energy detector may be based on the contribution of two or more components. Whilst this may be accounted for by a models that extract velocity on the basis of, for example, a regression in frequency space (Heeger, 1987), there are a number of motion patterns, the perceived velocity of which cannot be accounted for in these terms. Such patterns include the classes of motion stimuli defined as *second-order* or *non-Fourier.*

1.3. The detection of second-order motion.

The classes of second-order and non-Fourier motion stimuli contain stimuli whose perceived motion cannot be readily explained in terms of orientations of Fourier motion energy through the frequency domain origin. Although these classes of motion are defined differently, the term "second-order" has generally been applied to both. Except where necessary, both of these classes of stimuli shall be referred to as second-order.

1.3.1. Second-order motion.

The first-order/second-order dichotomy was first suggested by Cavanagh & Mather (1989). They defined first-order characteristics as those characteristics that are encoded at the lowest level in vision, at the level of retinal ganglion cells; ie. luminance and colour. As the M pathway has been identified as the motion pathway, and as this pathway is generally seen not to encode colour information, there has been a considerable reluctance to accept the proposal that colour-defined motion and luminance defined motion are processed by the same mechanism. Research into the perception of motion elicited by isoluminant stimuli suggests that luminance motion and chromatic motion are either processed differently or by different mechanisms (for a short review see Gegenfurtner & Hawken, 1996a). Therefore, more generally, the term first-order motion has come to mean rigid luminance defined motion. As detailed above, this can be characterised by orientations of Fourier motion energy through the frequency domain origin (Fleet & Langley, 1994).

In the definition provided by Cavanagh & Mather, second-order motion is defined by the translation of second-order statistics. The second-order statistics of an image are defined as the frequency with which particular combinations of luminance or colour values occur. The class of second-order motion stimuli therefore include the motion of contrast modulations, velocity modulations, disparity modulations and texture modulations. Some of the most common stimuli of this type are contrast modulations, specifically the motion of contrast modulated sine waves and beats created from the addition of two sine waves. Examples of these stimuli and their associated Fourier spectra are shown in Figures 1.6 and 1.7 respectively. What is particularly interesting about these stimuli is that under most conditions, the motion perceived by subjects is given by orientations of Fourier energy that do not lie through the frequency domain origin (Fleet & Langley, 1994). This is indicated by the dotted lines in the Fourier spectra. The motion that may be elicited by these stimuli cannot be accounted for by the Fourier components approach described above.

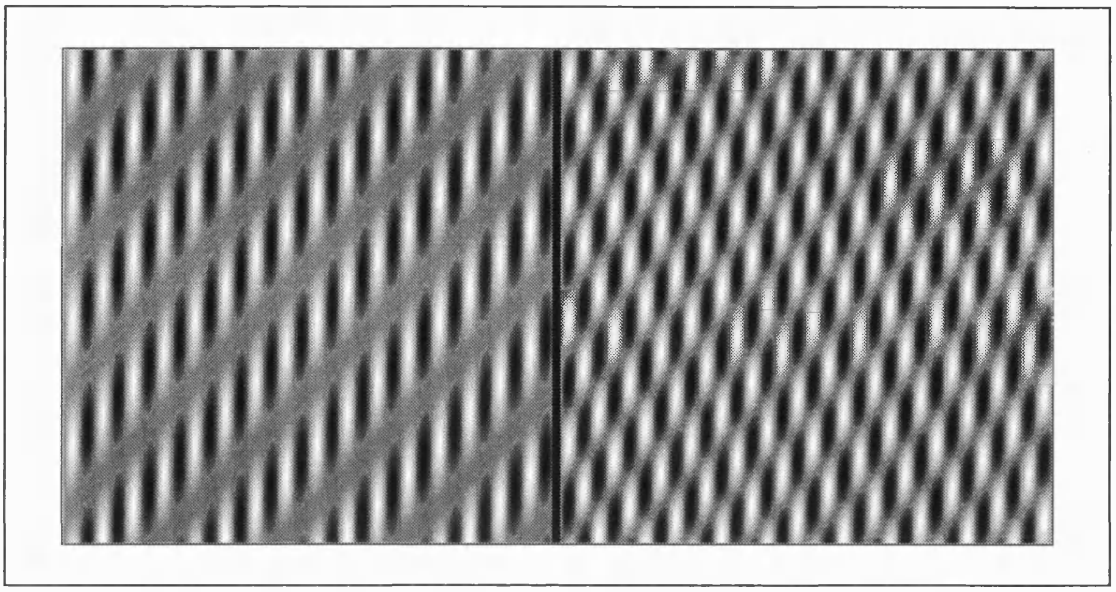


Figure 1.6: Space-time plots of contrast modulated sine wave grating (left) and a beat pattern created by the addition of two sine waves (right). Images shown are space time plots with increasing time as one moves down the vertical axis and increasing displacement as one moves along the horizontal axis.

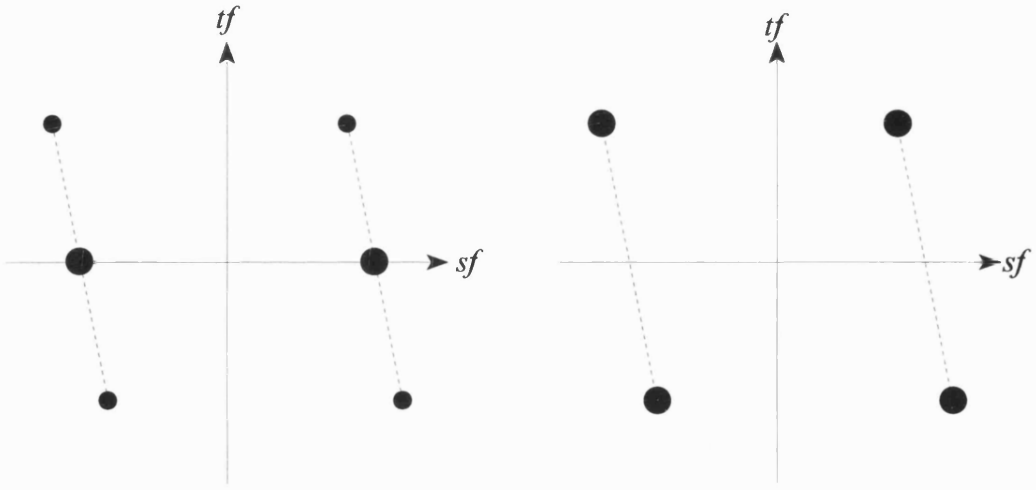


Figure 1.7: Fourier spectra of contrast modulated sine wave grating (left) and beat pattern created by the addition of two sine waves (right). The areas of the dots marking the components are approximately proportional to the amplitudes of the components.

1.3.2. Non-Fourier motion.

Chubb & Sperling (1988) describe a number of stimuli, the motion of which cannot be extracted using an approach to motion perception that is based on Fourier decomposition. They introduce the terms *drift-balanced* and *micro-balanced.* These are defined as follows:

"A random stimulus *S* is drift balanced if its expected power in the frequency domain is symmetric with respect to temporal frequency",

and

"*S* is microbalanced if the result *WS* of windowing *S* by any space-time-separable function W is drift balanced". (Chubb & Sperling, 1988, p1986).

The class of microbalanced stimuli includes contrast and frequency modulations of binary noise. Space-time plots and luminance profiles of these stimuli are shown in Figure 1.8. Chubb & Sperling provide a formal mathematical proof showing that these stimuli are indeed microbalanced. By their definition such a stimulus cannot consistently elicit Fourier motion. It is for this reason that the motion elicited by these stimuli has been termed *non-Fourier.* It should be noted that the Chubb & Sperling definition does not imply that no Fourier motion can ever be detected in these stimuli. It is the fact that no *consistent* Fourier motion can be detected in these stimuli. As it says in the definition, at any spatio-temporal location, the *expected* power in the frequency domain is symmetric with respect to temporal frequency. Therefore the actual power may not be symmetric but the probability of that power signalling motion in the leftwards direction is equal to the probability of it signalling motion in the rightwards direction.

The non-Fourier motion in the stimuli described by Chubb & Sperling is readily perceived by human observers. Therefore the principle of deriving motion from Fourier components cannot fully account for human motion perception. This presents us with the following three options. Firstly, that the Fourier energy mechanisms need to be adjusted so that they become sensitive to second-order motion. Secondly, that some additional non-Fourier mechanism exists that is sensitive to second-order motion. Finally, that the Fourier energy approach should be abandoned. The first and second of these options have been the most popular. The Fourier motion energy approach has been extended to account for second-order motion

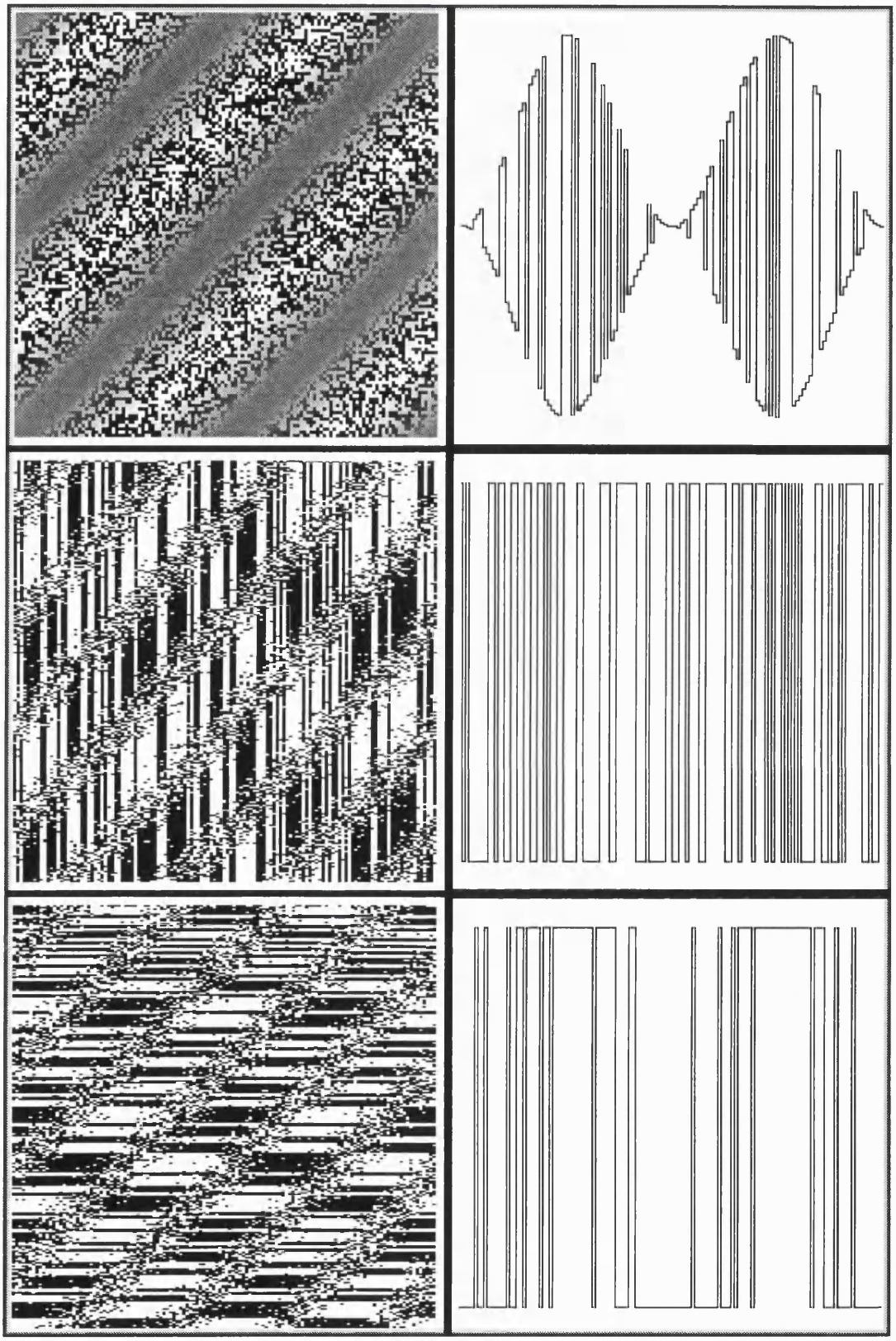


Figure 1.8: Microbalanced stimuli, left hand side shows space-time images, right hand side shows associated spatial cross sections of luminance profiles. Shows (top) contrast modulated binary dynamic noise, (middle) temporal frequency modulations, and (bottom) spatial frequency modulation. In each space-time plot, time increases as one moves down the vertical axis, displacement increases as one moves along the horizontal axis.

through the distortion product hypothesis. The addition of a non-Fourier channel has generated a two channel model of human motion processing.

1.3.3. The distortion product hypothesis.

Burton (1973) found that a beat pattern created by the addition of two high contrast interference fringes was readily distinguishable. Each of the interference fringes was of a high enough spatial frequency to be unresolvable. Burton suggested that a logarithmic non-linearity inherent in the visual transduction process may introduce distortion products into the signal at the spatial frequency of the beat. This idea may be extended to the perception of second-order motion. Any one of a number of non-linearities can introduce distortion products at the frequency of the second-order components into the signal. In this approach, both the first- and second-order motion information is carried in one single channel. The same motion detectors that extract first-order motion will also serve to extract second-order motion.

Henning, Hertz & Broadbent (1975) examined masking effects between sine wave gratings and complex gratings. The latter were constructed from the addition of three sine wave gratings of the same orientation but differing spatial frequencies such that they were mathematically equivalent to the contrast modulation of a high frequency carrier with a low frequency envelope. Henning *et al* found that masking using sine wave gratings of the same spatial frequency as the envelope significantly reduced the detectability of the complex grating. There was also a strong effect on the detectability of envelope frequency sine wave gratings when these were masked with complex gratings. When the phase relationship of the three components comprising the complex grating were changed so that the variation of contrast across the pattern was significantly reduced, there was no strong masking effect. These results are consistent with the idea of a distortion product.

However, Henning *et al* (1975) put forward the hypothesis that elements are sensitive to both low spatial frequency luminance modulations and low spatial frequency contrast modulations of high spatial frequency carriers. The notion that a distortion product may be the underlying cause of the pattern of results was rejected because there was no apparent interaction between a high contrast sine wave mask and a signal at the second harmonic of

that mask. Henning *et al* (1975) do make the point that this latter finding could be accounted for supposing that a non-linearity was applied to high frequency, but not low frequency, signals. Such a process would produce distortion products at the frequency of the modulant with a high frequency carrier but would not produce any distortion products with a low frequency sine wave mask. This might reasonably account for their results and those of Burton (1973).

Nachmias & Rogowitz (1983) used a similar stimulus and replicated the findings of Henning *et al* (1975). They also showed that, with a contrast modulated sine wave mask, the strength of the masking effect was strongly phase dependent. They tentatively hypothesise that only a distortion product at the spatial frequency of the envelope caused by some non-linearity in the visual system can account for the interactions between contrast modulated and sine wave gratings. However, they do qualify this by noting that some of the results found in their study do not agree with this hypothesis. Particularly, they find discrepancies in the relationship between the estimated size of the distortion product and the contrast and the consistency of the complex pattern. It would seem from the results of Henning *et al* (1975) and Nachmias & Rogowitz (1983), that there is some interaction between modulations of contrast and sine wave gratings but that the actual way that these may interact is not clear.

Derrington & Badcock (1986) examined the detection of beats created by the addition of two sine waves. With such a pattern it can be shown that the amplitude of the distortion product at the beat frequency is proportional to the product of the amplitude of the component sine waves. Therefore, if the amplitude of one of the components is increased then the amplitude of the other would need to decrease to maintain the same level of detectability. However, increasing the amplitude of one of the components also has the effect of reducing the difference in contrast between the high and low contrast regions in the beat pattern. With this arrangement Derrington & Badcock attempted to test whether the perception of the beat was based on comparative contrast or a distortion product. If detection of beats is based on comparative contrast then, if the amplitude of one of the components is increased, the amplitude of the other component also needs to be increased to maintain the same level of detectability. Derrington & Badcock found this to be the case.

In a direct test of the distortion product hypothesis, Badcock & Derrington (1989) showed that the perceived motion of a drifting beat could not be nulled by the addition of a signal designed to cancel the distortion product. Evidence against the distortion product hypothesis is also provided by the following studies. Hammett & Smith (1994) examined temporal beats and found no masking effects. A luminance non-linearity should produce distortion products as the signal changes over time and should therefore affect the detectability of a grating that is temporally modulated at the beat frequency. Ledgeway & Smith (1994a) interleaved first- and second-order motion gratings and showed that, when there was a quarter cycle displacement between each frame, there was no consistent percept of motion direction. If an expansive non-linearity is applied to a contrast modulated pattern then a distortion with the same frequency and phase as the modulant is introduced into the signal. If this were the case then the stimulus employed by Ledgeway & Smith should appear to move forwards. If a compressive non-linearity is applied then the distortion product will be π radians out of phase with the modulant. The stimulus should appear to move in the direction opposite to the displacement. Therefore the Ledgeway & Smith (1994a) finding appears to weigh against the distortion product hypothesis. A similar study by Mather & West (1993) concurs with the findings of Ledgeway & Smith (1994a).

The evidence outlined above implies that the distortion product hypothesis cannot account for the perception of second-order motion. Recently the picture has become more complex. Holliday & Anderson (1994) obtained a bimodal contrast sensitivity function for direction discrimination in beats as a function of envelope temporal frequency. They subsequently examined the effects of adaptation and cross-adaptation on direction discrimination thresholds between the four possible permutations of beat and luminance grating stimuli over temporal frequency. At high beat envelope and luminance temporal frequencies the shapes of the four functions appeared identical. This was not the case at low temporal frequencies where little or no cross-adaptation was observed. Holliday & Anderson conclude that this pattern of results implies that, at high envelope temporal frequencies, the luminance and beat patterns are detected by a single mechanism and that the detection of the second-order signal is consistent with the notion of a distortion product. A similar conclusion has also been reached by Scott-Samuel & Georgeson (1995) (see also Scott-Samuel, 1997). At high envelope temporal frequencies, they successfully managed to null the motion elicited by a second-order motion sequence by adding a luminance modulation

to the pattern. Additionally, the finding of Ledgeway & Smith (1994a) which showed no consistent motion elicited by an interleaved first- and second-order gratings stimulus was replicated but only at a low temporal frequency. At a high temporal frequency, coherent motion was elicited by this stimulus in the direction opposite to the displacement; a result that would be predicted by the operation of a compressive non-linearity.

Although there is evidence for a role for distortion products at high envelope temporal frequencies, there is no strong evidence to suppose that they can account for second-order motion at low temporal frequencies. If this is the case then some non-Fourier mechanism must exist. However the idea that second-order motion is extracted via some non-linearity that precedes a motion detection stage has not been abandoned. Chubb & Sperling (1988) suggest that a specific non-linear mechanism or channel exists for the extraction of second-order motion and that this channel exists in parallel to the traditional idea of the Fourier motion energy channel (Chubb & Sperling, 1989a).

1.3.4. The two-channel model.

Chubb & Sperling (1988), to account for the perception of motion in the class of microbalanced stimuli put forward their "non-linear transformations hypothesis". They suggest that some gross non-linearity such as fullwave rectification may precede standard frequency domain power analysis. Non-linearities of the kind discussed by Chubb & Sperling are far more effective at recovering second-order motion than, for example, the logarithmic non-linearity suggested by Burton (1973). In terms of frequency space, the effect of rectification is to collapse those orientations of Fourier energy that describe the second-order motion onto a line that lies through the frequency domain origin. A space-time plot of a fullwave rectified contrast modulated sine wave grating is shown in Figure 1.9 (top), the Fourier spectrum of the rectified image is shown in Figure 1.9 (bottom). Rectification makes the second-order motion signal recoverable by a simple energy mechanisms of the kind described by Adelson & Bergen (1985). Chubb & Sperling (1988) propose that the non-linearity is preceded by spatio-temporal filtering. This stage is necessary to explain the perception of motion in spatial and temporal frequency modulations of binary noise. Chubb & Sperling (1989a) describe a reverse-phi stimulus that consists of a series of bars that step forward by a quarter of a cycle every frame and reverse their contrast polarity. At short

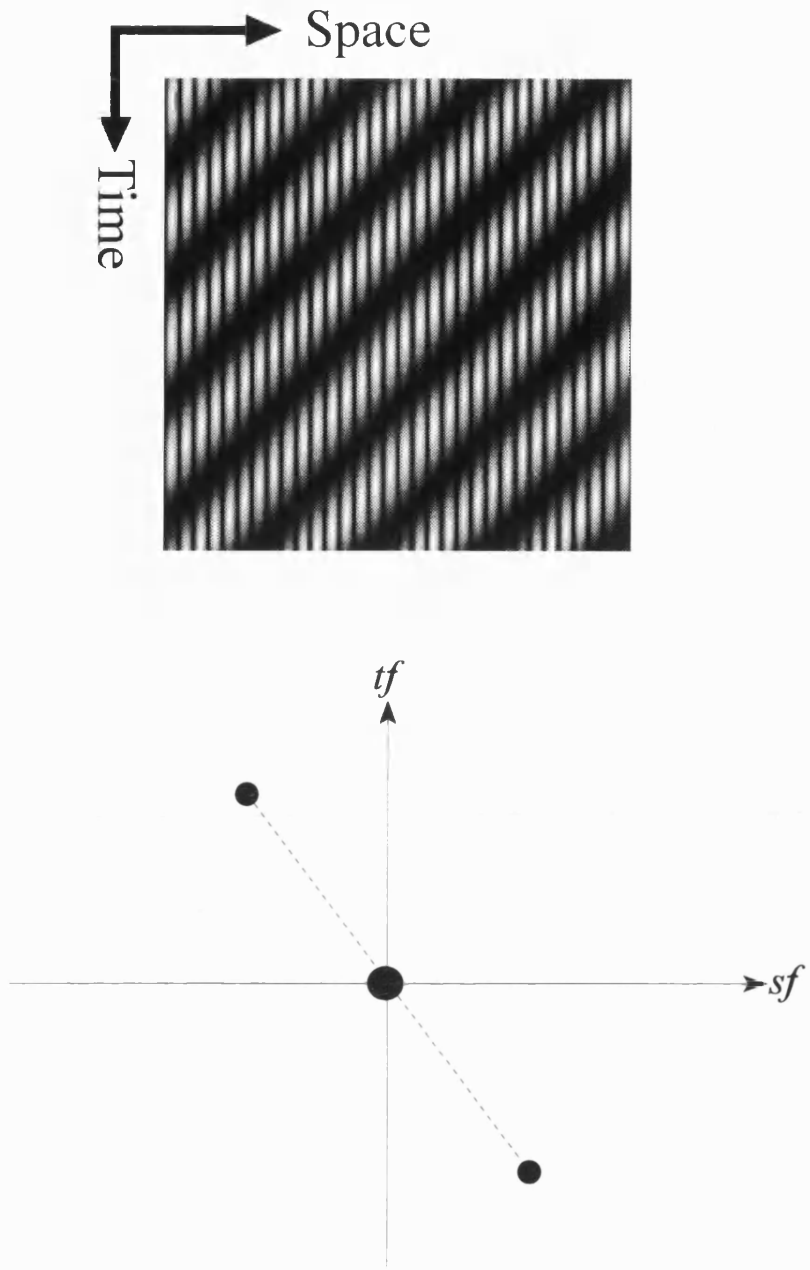


Figure 1.9: (top) Space-time plot showing rectified contrast modulated sine wave grating, and (bottom) Fourier spectrum of the rectified image. The areas of the dots are approximately proportional to the amplitude of the components.

viewing distances this stimulus appears to moving forwards. As viewing distance is increased, the stimulus appears to be going in the reverse direction. Chubb & Sperling take this indicate the existence of two separate processing streams, a first-order channel and a second-order channel. Such a model has also been proposed by Wilson, Ferrera & Yo (1992). A schematic diagram of the two channel model is shown in Figure 1.10.

The Chubb & Sperling (1989a) contrast reversing stimulus is particularly interesting. In terms of the application of a pointwise non-linearity, some fullwave rectification process (or similar) applied around mean luminance would be necessary to make motion in the direction of displacement visible to standard motion energy analysis. The motion in this stimulus cannot be explained through the operations of some monotonically increasing function applied directly to the luminance input. In other words, the distortion product hypothesis cannot account for the perception of forward motion that is readily elicited by this stimulus.

Chubb & Sperling (1989b, 1991) have examined the form that the initial linear filtering process may take. They showed that, to account for the perception of motion in both a contrast modulating square-wave and a contrast reversing square wave, the initial temporal filtering stage must include filters with both lowpass and bandpass characteristics. They postulated a "best of both worlds" temporal filter which is constructed by the addition of these two types of filters. The spatial filtering characteristics were examined using texture quilts. Chubb & Sperling (1991) showed that subjects could readily perceive spatial frequency and texture orientation modulations and that there seemed to be no consistent difference in performance between these two type of stimuli. They put forward the proposal that a stage of oriented bandpass (specifically Gabor) filtering prior to some point-wise non-linearity might reasonably account for these data. This view is also supported by Langley, Fleet & Hibbard (1996), Wilson *et al* (1992) and Sutter, Sperling & Chubb (1995).

At present there is no firm consensus on the form of the subsequent non-linearity. Sperling (1989) and Lu & Sperling (1995a) appear to favour fullwave rectification, a process that could be accomplished by the simple addition of on and off pathways whilst Wilson *et al* (1992) suggests squaring as the non-linear process. Chubb, McGowan, Sperling & Werkhoven (1994) provide evidence which supports the existence of two second-order motion channels. They argue that the "second" second-order channel uses a thresholded

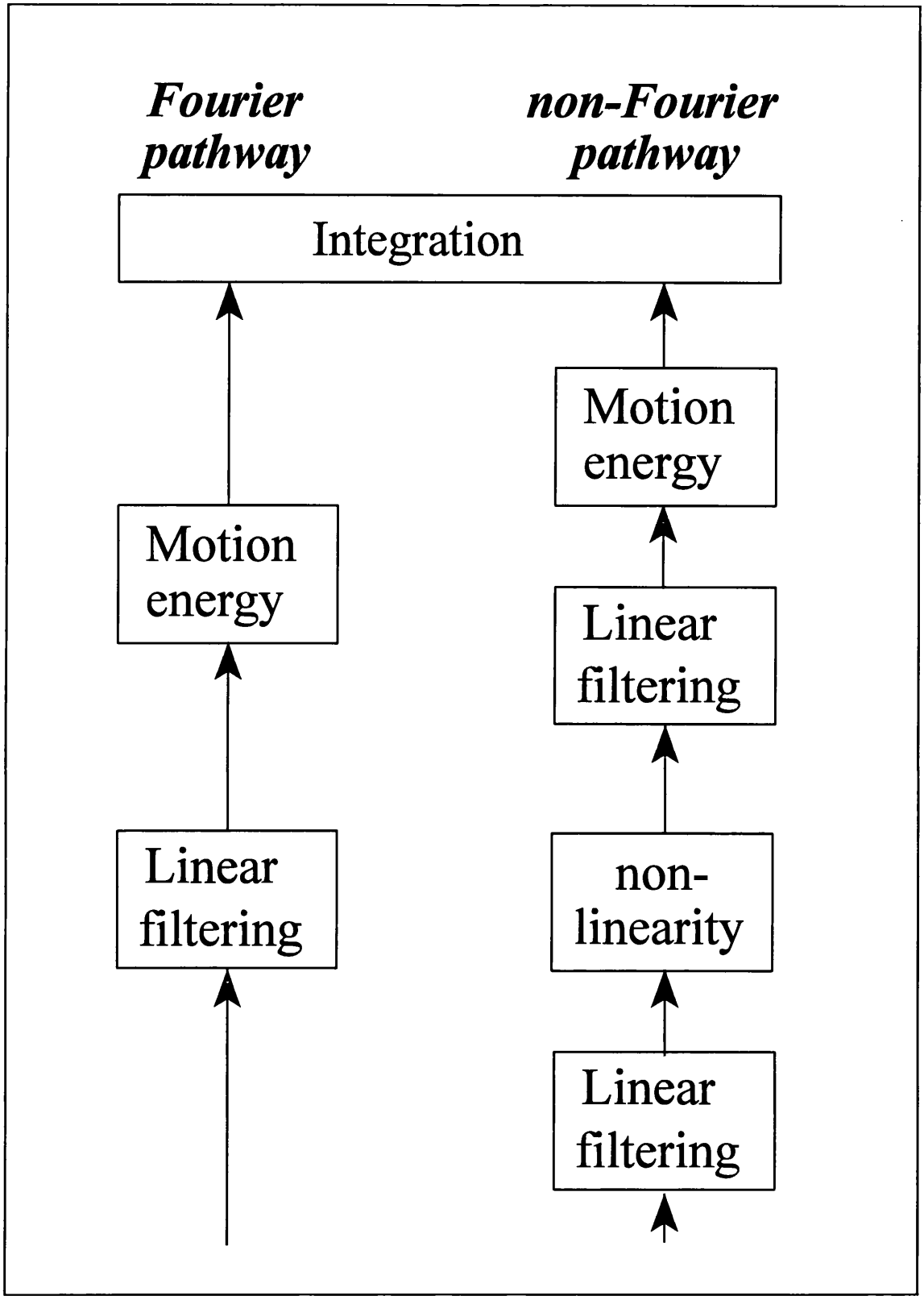


Figure 1.10: Schematic diagram of two channel model outlined by Wilson *et al* (1992).

pointwise non-linearity. Solomon & Sperling (1994) employed second-order stimuli that they argue could only be perceived using halfwave and not fullwave rectification (and vice versa). They conclude that there are separate fullwave and halfwave channels but that only approximately one third of the population can perceive stimuli designed for the latter. It is probably unrealistic to expect a biological system to be perfectly linear and, as we have seen, it has also been suggested that there is a distortion product within the first-order channel as well as an explicit non-Fourier mechanism (Holliday & Anderson, 1994; Scott-Samuel & Georgeson, 1995).

Wilson *et al* (1992) have proposed a specific mechanism for the integration of first- and second-order signals using a lateral inhibition competitive neural network to account for the perception of motion direction in plaids. Contained within this model is the idea that motion detection in first- and second-order signals relies on different sets of detectors and that the information from these is subsequently combined. There is some psychophysical evidence that challenges this assumption. Ledgeway & Smith (1994b) found cross-adaptation between first-order adapting stimuli and second-order test stimuli (and vice versa). In fact, they found no evidence of any difference between cross-adaptation and same-adaptation conditions. This finding might imply that both types of motion are detected by a common low level motion detection mechanism. Turano (1991) found that cross adaptation between sine wave gratings and contrast modulated sine wave gratings (and vice versa) occurred, she concludes that this supports the idea of a common low level mechanism. This conclusion is also supported by the following two observations: Firstly, velocity discrimination thresholds in sine wave gratings and contrast modulated sine waves appear to be very similar both in magnitude and in the way that they vary with velocity (Turano & Pantle, 1989). Secondly, perceived velocity in both contrast and luminance defined stimuli varies similarly with modulation depth (Ledgeway & Smith, 1995).

We can summarize the two channel model in the following manner. There exists a luminance (Fourier) channel which appears to contain some small intrinsic distortion product. There is at least one grossly non-linear second-order channel based on a fullwave rectification-like process. In this pathway the non-linearity is preceded by oriented bandpass spatial filtering, with filters tuned to a range of spatial frequencies, and a mixture of both bandpass and lowpass temporal filtering. After the non-linearity, the second-order system is tuned to lower

spatial frequencies than the first-order system (Sutter, Sperling & Chubb, 1995; Wilson, 1994).

The construct of a second-order channel is largely based on the implicit acceptance of the principle that motion can only be detected from Fourier components that lie on a plane or line through the frequency domain origin. If second-order motion cannot be fully explained by the distortion product hypothesis then, it is argued, the only way that motion can be extracted is by constructing a non-linear channel designed specifically to ensure a signal with this property. However, Johnston & Clifford (1995a) showed that it was possible to account for the perception of second-order motion in contrast modulated gratings using a model without an explicit non-linear stage preceding motion detection. They have also demonstrated that the reversal of direction found by Chubb & Sperling (1989a) can be accounted for without appealing to a separate non-linear channel (Johnston & Clifford, 1995b). The simple fact that the Johnston *et al* (1992) model can extract second-order motion has interesting ramifications. Firstly, it implies that a non-linearity that precedes motion calculation is not necessary in order to model the perception of second-order motion. Secondly, it implies that the behaviour of the model utilised by Johnston & Clifford cannot be accounted for simply in terms of the Fourier decomposition approach.

1.3.5. Evidence for the two channel model.

Good reviews of possible differences in processing between first- and second-order motion are provided by Ledgeway (1994a), Smith (1994a) and Bruce *et al* (1996). Whilst there are differences there are also substantial similarities between the patterns of response to the two types of stimulus. Smith suggests that the close similarities imply that common processing principles (ie. motion energy detection) whilst the differences imply different mechanisms. The overview of research presented below is merely a portion of a the huge body of research on second-order motion perception. To maintain some degree of coherence I have mainly described the findings from two strands of research carried out on the one hand by Derrington and his colleagues and on the other hand by Smith and his colleagues (see below for references). A great deal of research into second-order motion perception is also described in the two preceding sections.

Derrington, Badcock & Henning (1993) examined the detection of motion direction in a luminance grating and a beat pattern. They showed that the motion of the beat could only be discriminated at stimulus durations of more than 200 msec. The direction of motion in the luminance grating was discriminable at far lower stimulus durations. Derrington *et al* take this as evidence for the existence of a second-order mechanism which operates more slowly than the first-order mechanism. In agreement with this finding, Cropper & Derrington (1993) measured the lowest velocity required to detect motion in luminance gratings and beats presented at multiples of their respective detection thresholds. Direction of motion was not discriminable in beat patters below stimulus durations of 120 msec whereas the motion in luminance gratings was discriminable at stimulus durations as low as 15 msec (see also Cropper & Derrington, 1996). Evidence for a difference in processing speed between first- and second-order stimuli is also supplied by Yo & Wilson (1992) who found that the perceived direction of motion in plaids could vary as a function of stimulus duration. Wilson *et al* (1992) found that this effect could be modelled by incorporating a short delay into the second-order channel of their two channel model.

Not only does it appear that the motion of beats is perceived more slowly than the motion of luminance gratings, there is also evidence that the function describing contrast thresholds for detecting motion direction in beats is tuned to lower temporal frequencies than the function describing thresholds for the detection of motion direction in luminance gratings (Derrington & Henning, 1994). However, more recently, Lu & Sperling (1995b) showed that the temporal contrast sensitivity functions for direction discrimination were similar for a luminance modulation and a contrast modulation of static binary noise. Similar evidence has also been produced by Smith & Ledgeway (1997b) who additionally showed that, whilst the contrast sensitivity function for direction discrimination with modulations of dynamic noise appeared lowpass, those for luminance modulations and modulations of static noise appeared bandpass. These data imply that the form of the tuning function for second-order stimuli is dependent upon the particular stimulus that is being investigated.

Recently a programme of research carrier out by Ledgeway and Smith (Ledgeway & Smith, 1994a; Ledgeway & Smith, 1994b; Ledgeway, 1994b; Ledgeway & Smith, 1995; Ledgeway & Smith, 1996) has attempted to identify differences in processing between first- and second-order motion. This series of studies has the advantage that they are carried out using

microbalanced stimuli rather than, for example, contrast modulated gratings or beats. Direction of motion in contrast modulated gratings can potentially be detected by an energy model (Chubb & Sperling, 1988). Ledgeway and Smith added noise to their first-order stimuli so that any differences between the two types of stimuli could not be put down to the presence (or absence) of the carrier. They also took the precaution of equating the visibility of the two types of stimulus by presenting them at the same multiple of their respective direction identification thresholds. In all of their studies comparing first- and second-order motion no reliable differences were found to discriminate measures of performance in response to the stimuli.

For example, Ledgeway & Smith (1994b) and Ledgeway (1994b) examined the motion aftereffect for first- and second-order motion. They found both first- and second-order motion produced similar aftereffects and that cross-adaptation between the two classes of stimuli also occurred. Müller & Greenlee (1994) showed that the perceived speed of a luminance grating could be reduced by adaption to a grating of the same spatial frequency and of the same or a higher temporal frequency. Similarly, Ledgeway & Smith (1996) examined the effects of adaptation and cross-adaptation to first- and second-order motion on the subsequent perceived speed of first- and second-order motion sequences. They showed that prior adaption to both types of grating could affect the perceived speed of both first- and second-order stimuli. All four adaption/cross-adaptation conditions showed similar patterns of response to one another. When the speed of the adaptation stimulus was higher than that of the subsequently presented test stimulus, a reduction in perceived speed was observed. However when the adaptation stimulus moved more slowly than the test stimulus then increases in perceived velocity were observed. Ledgeway and Smith proposed that the data supports two hypotheses. Firstly, that both types of motion are detected by the same mechanism and secondly, that they are detected by different mechanisms operating on the same principle. In the papers described in this paragraph, Ledgeway and Smith come down on the side of the latter hypothesis.

A number of studies have shown that the contrast of a stimulus affects its perceived velocity (Thompson, 1982; Stone & Thompson, 1992; Thompson, Stone & Swash, 1996). Ledgeway & Smith (1995) examined the perceived speed of second-order motion. When equated for visibility, the perceived speed of the first- and second-order stimuli was the

same. They also looked at the effect of varying modulation depth on the perceived velocity of first and second-order stimuli. In agreement with studies examining the effect of reducing contrast on first-order stimuli, Ledgeway & Smith found that the perceived speed was linearly related to the log of the modulation depth. Interestingly, the slope of this line (determined by linear regression) was no different for the two stimulus types. Gegenfurtner & Hawken (1996b) also examined the slope of the line relating log contrast to the log of perceived velocity and found that the slope of the line was substantially greater for second-order stimuli (square wave modulations of binary noise) than for first-order stimuli. Gegenfurtner & Hawken used sine wave gratings as their first-order stimuli and amplitude modulated binary noise as their second-order stimuli. It is possible that the difference between the two studies can be directly attributed to the additive noise used by Ledgeway & Smith in their first-order stimuli.

Edwards & Badcock (1995) studied the perception of motion in stimuli that consisted of mixtures of Fourier and non-Fourier dots. They measured motion thresholds in dot stimuli by varying the number of signal dots within a fixed number of dots (the remainder would move randomly). With the addition of random motion Fourier dots, motion thresholds for non-Fourier dots decreased substantially but the addition of random motion non-Fourier dots had little effect on the motion threshold for the Fourier dots. In a second experiment, Edwards & Badcock looked at global motion thresholds in stimuli consisting of the two dot types where a proportion of the dots of each dot type moved in the opposite direction to a proportion of those of the other type. The data from this experiment supported that of the first in that, to reach motion threshold, the coherence of the second-order dots needed to be substantially increased. This was not the case with the first-order dots. Edwards & Badcock take these data to indicate the existence of two channels; a first-order channel that is sensitive to the motion of the Fourier dots and a second-order channel that is sensitive to the motion of both the Fourier and the non-Fourier dots. This study is of particular interest because it does appear to show a genuine difference between the processing of first- and second-order signals.

In a recent study, McOwan & Johnston (1996a) examined the perception of transparency elicited by stimuli containing translating Fourier dots or translating non-Fourier dots. In a condition where transparency was consistently perceived by a number of subjects, replacing

the Fourier dots with non-Fourier dots in the display led to the abolition of transparency. When the consistency of the non-Fourier dots was unaltered, but the background luminance was changed so that a first-order signal was introduced, then the perception of transparency was reinstated. Initially this would appear to be a clear indication of a difference between first- and second-order processing with the evidence supporting the theory that the second-order system had no capacity to support transparency. However, when dynamic noise was added to the background in the Fourier dot stimulus, then the percept of transparency was abolished. This latter finding would seem to imply that noise has some critical part to play in the perception of transparency and McOwan & Johnston conclude that lack of transparency in the stimulus containing the non-Fourier dots is based on the presence of noise in the local velocity field.

In one study, Ledgeway & Smith (1994a) examined the perception of motion in a stimulus created by interleaving first- and second-order patterns. When the patterns step forwards a quarter cycle every frame and alternate between first- and second-order then no motion is seen in this stimulus. Ledgeway & Smith argue that if both first- and second-order motion were detected by a single system then the information from both types of pattern would be integrated to form an unambiguous motion percept. They take this as evidence for strong proof for the two channel hypothesis. A similar study by Mather & West (1993) produced similar findings and conclusions. However, Johnston & McOwan (1996) recently demonstrated that this finding could be accounted for by a single channel model. Furthermore, Scott-Samuel (1997) showed that motion could be elicited by the Ledgeway & Smith stimulus with increases in temporal frequency. An additional piece of strong evidence for the two channel model is the finding that second-order motion does not elicit optokinetic nystagmus (OKN) (Harris & Smith, 1992). This finding is weakened by the fact that Harris & Smith used a standard drifting sine wave grating as their first-order stimulus. They cannot guarantee that the presence of the carrier in the second-order stimulus might be responsible for the low level of OKN.

Studies examining human patients have failed to provide reliable evidence of cortical differences between the processing of first- and second-order motion. Vaina & Cowey (1996) looked at the perception of first- and second-order motion in a subject with a small unilateral lesion adjacent to human MT/V5. They find some performance deficit with two

types of second-order motion when compared to performance on a first-order motion stimulus. However, in a more recent study, Greenlee & Smith (1997) examined orientation, direction and speed discrimination thresholds in first- and second-order motion patterns. They studied these in 21 patients with unilateral damage to various regions of cortex. By comparing the responses of the patients over their hemifields and by taking into account the different regions of damage, Greenlee & Smith hoped to be able to localise an area of the brain that might be implicated in second-order motion. Whilst they do find slight quantitative differences between the two stimuli, they find no real evidence for anatomical specialisation between first- and second-order motion processing. Furthermore, in a study which recorded visual evoked potentials (VEPs) in response to Fourier and non-Fourier motion, Victor & Conte (1992) found no evidence for any differences in response either in temporal onset or scalp topography of the components of the VEP that were examined. Using MRI, Greenlee, Smith, Singh, Kraemer & Hennig (1997) attempted to locate the cortical areas involved in the perception of first- and second-order motion. They found no difference in the activated areas between first-order motion and three types of second-order motion; contrast modulated static noise, contrast modulated dynamic noise and flicker frequency modulated noise. Greenlee *et al* reported that second-order motion readily activated an area identified as human V5/V5a.

The evidence from recording responses of cells in the visual pathways of cats and primates are similarly inconclusive. It has been demonstrated that cells in area MT of primates can respond to drift balanced second-order motion (Albright, 1992). Of course, this says nothing about how these cells come to respond to the second-order motion. In cats, there is evidence that cells in areas 17 and 18 respond to envelope motion in contrast modulated gratings (Zhou & Baker, 1993, 1994; Mareschal & Baker, 1997). There is also evidence that cells in area 18 of cat respond to motion in non-Fourier stimuli (Mareschal & Baker, 1996). Another study examining responses of cells in the primary visual cortex of both cats and primates (Albrecht & DeValois, 1981) found no evidence of a response to drifting contrast modulated gratings. It is difficult to say what pattern emerges from these data. That we should obtain some response to second-order motion is unsurprising, after all we can see second-order motion (we assume that other primates and cats also have this facility). The envelope responsive cells studied by Zhou & Baker responded to both first- and second-order motion. The relationship between the spatial frequency tuning for first-order motion,

and the spatial frequency tuning for envelope motion rules out the possibility of explaining the result by any early pointwise non-linearity. To account for their data, Zhou & Baker propose that their envelope responsive cells receive input from both linear and non-linear pathways. Of course, all we really know about these cells is that they respond to both first- and second-order motion. The mechanism that Zhou and Baker propose is pure hypothesis. It is interesting to note that nobody has found cells that respond to second-order motion whilst giving little or no response to first-order motion. Such cells would be predicted by the two-channel model.

In general, there appears to be no conclusive evidence for a difference in processing between first- and second-order stimuli. Where differences have been found, subsequent investigation has not supported the distinction. For example Pantle (1992) found the motion of some peripherally presented second-order stimuli was not readily distinguished. However, Solomon & Sperling (1994) did observe sensitivity to second-order motion in the periphery. The difference between the two studies appears to be due to the fact that Solomon & Sperling (1994) utilised much larger patches of stimulus than Pantle (1992). Similarly a conclusion that second-order motion could not support a significant motion aftereffect (Derrington & Badcock, 1985) was subsequently shown to be false (Ledgeway & Smith, 1994b). From the studies presented above there is also the implication that differences in the perception of motion elicited by first- and second-order motion stimuli may be based on the presence of the carrier (or some other difference) rather upon the "orderness" of the stimuli.

1.3.6. Feature tracking.

Whilst the low level detection of motion is thought to be based upon the extraction of motion by combining the results of local spatio-temporal filtering, there is another vein of research that has been taken to suggest the existence of a higher level feature tracking mechanism (Georgeson & Harris, 1990; Cavanagh, 1992; Hammett, Ledgeway & Smith, 1993). This has been identified with some aspects of the long range process identified by Braddick (1974, 1980). Feature tracking mechanisms may be divided into two broad types; those that track low-level features and those that track highly complex features. Examples of the latter are mainly applied in the field of machine vision. These mechanisms have been largely knowledge based and have been used to track body parts or features for the

recognition of biological motion (Goddard, 1989), gesture recognition (Davis & Shah, 1994), and gaze tracking (Gee & Cipolla, 1994). Low level feature tracking mechanisms combine the outputs of linear filters tuned to different frequencies or orientations (Watt & Morgan, 1985; Georgeson, 1992; Bowns, 1996). These early operations are generally seen as part of a process that detects edges and are generally followed by some procedure to extract zero crossings following the approach described by Marr & Hildreth (1980).

In many ways, the extraction of features through some low level processing procedure is similar to operation of second-order motion mechanisms. These might usefully be regarded as a feature tracking mechanisms that is specialised to extract second-order motion. Watt & Morgan (1985) present a model for the extraction of low level features that contains characteristics of both traditional feature tracking mechanisms and second-order motion mechanisms. In the first stage, the on-centre outputs from a range of spatial frequency filters are combined into one output stream and the off-centre outputs from the same filters are combined to form another output stream. These output streams are then subject to a processing stage in which the amplitude and width of "centroids" are identified. These are essentially zero bounded regions of activity in each of the output streams. These centroids are then labelled according to a series of rules which define whether they describe bars, edges or areas of zero activity. The resultant "spatial primitives" may subsequently be subjected to some motion analysis procedure.

The Watt & Morgan (1985) model is particularly interesting because it provides a method for the extraction of features that contains many of the elements of a second-order motion extraction mechanisms. The combination of either the on-centre or off-centre outputs is a good biologically plausible method of applying a gross non-linearity (ie. halfwave rectification) to the signal. The fact that this follows spatial filtering and is then followed by some smoothing process might well produce similar results to the spatial filtering stages applied within the two channel model. It is therefore possible to see how the notions of a second-order channel and a feature extraction mechanism may be fused together. Indeed, if a spatial primitives extraction process preceded motion energy analysis, or some other motion detection mechanism, then a feature detection model might potentially account for both first- and second-order motion.

Earlier studies focused on providing evidence for the existence of some short range motion processing system in addition to some long range process (Braddick, 1974, 1980). More recently the existence of some short range low-level system has become widely accepted and the focus of the research has changed to provide evidence for the existence of a long range feature tracking mechanism. For example, Georgeson & Harris (1990) presented subjects with a missing-fundamental square-wave grating that translated by a quarter of a cycle of the spatial frequency of the fundamental. Traditionally, this stimulus has been used to demonstrate the existence of some low level motion energy process (Adelson & Bergen, 1985). The Fourier spectrum of the stimulus is shown in Figure 1.11a. The spatial and temporal frequency of the missing fundamental are marked by the crosses on the spectrum. The area of the dots representing the components is proportional to the amplitudes of the components. It can readily be seen that the largest component in the spectrum signals motion in the direction opposite to that of the displacement. A standard motion energy approach would predict that motion in the reversed direction should be perceived. However, Georgeson & Harris added a mean luminance interstimulus interval (ISI) between the quarter phase translations and found that, with increases in the temporal extent of the ISI, subjects increasingly reported that the stimulus elicited a perception of motion in the direction of the displacement. Georgeson & Harris take this pattern of results to indicate the operation of two motion systems, a short-range process and a long-range feature based motion process. Their conclusion is supported by the results from a study by Hammett *et al* (1993) who employed a beat stimulus created from the addition of two components with spatial frequencies of 3 and 4 times that of a notional fundamental frequency. They stepped the stimulus forward by a quarter of a cycle of the fundamental frequency, the Fourier spectrum of the stimulus is shown in Figure 1.11b. With no ISI, subjects reported that the stimulus appeared to exhibit motion transparency with both forwards and reversed motion present. As the ISI was increased, the percept of transparency decreased and subjects increasingly perceived motion in the direction of displacement. Hammett *et al* take this to indicate the existence of some long range feature processing mechanism.

In both of the studies described in the paragraph above it should be noted that the stimuli contained both first- and second-order components (see Figure 1.11). The orientation of the largest pair of components indicating second-order motion signals motion in the direction of translation. Both studies have been taken to indicate the existence of some feature

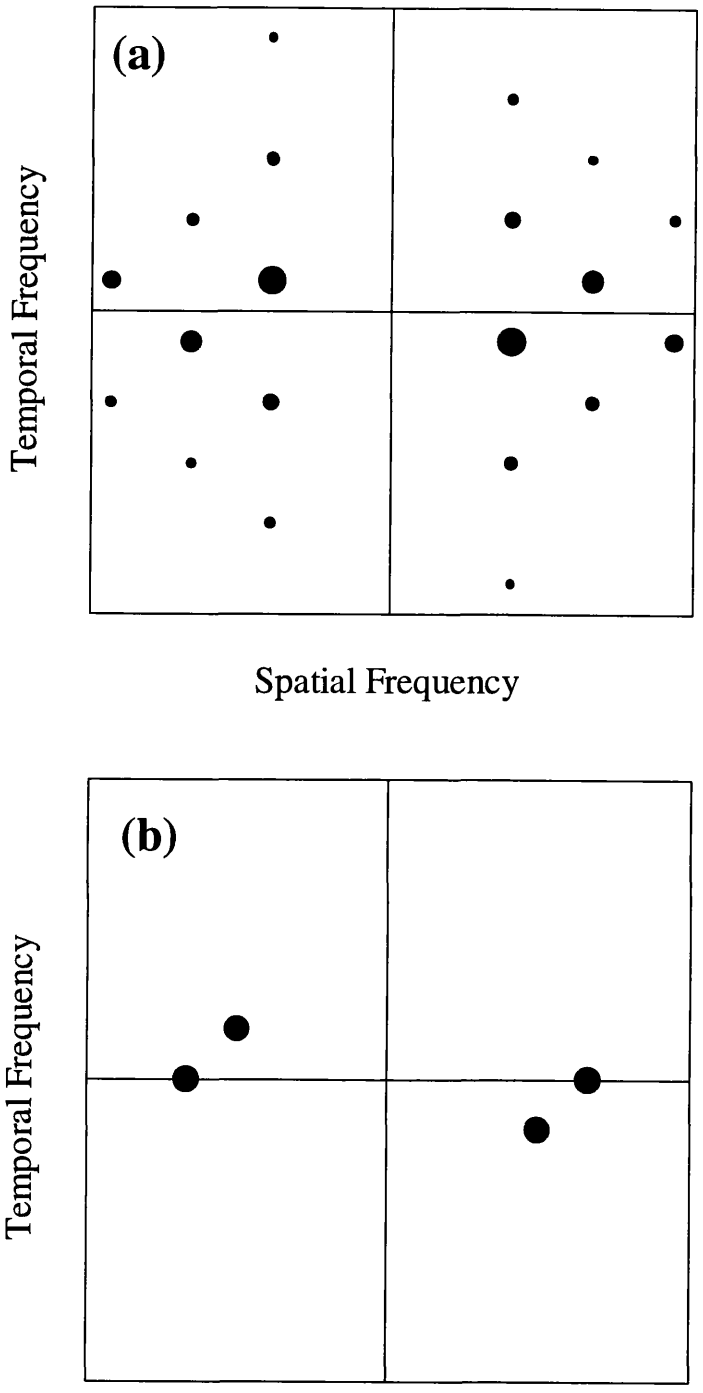


Figure 1.11: (a) Fourier spectrum of missing fundamental square wave grating that translates in steps that are equal to a quarter of a cycle of the missing fundamental. (b) Beat pattern created by addition of two components with spatial frequencies of $3f$ and $4f$ - the pattern translates in steps that are equal to the wavelength of the $4f$ component. In both spectra, the area of the dots marking the components is proportional to the amplitude of the components.

tracking process. As the ISI is increased, the contribution of the short-range process was thought to have decreased, leaving the long-range process increasingly dominant. However, the operation of some second-order mechanism (rather than a feature tracking mechanism) might also explain this pattern of results if it were supposed that the second-order mechanism were comparatively unaffected by increasing ISI. An additional possibility is that the patterns of results might be modelled by the operation of a single motion extraction mechanism. Johnston & Clifford (1995b), utilising an extended gradient model similar to that described by Johnston *et al* (1992), showed that the effect reported by Georgeson & Harris (1990) could be accounted for by the operation of a single low level motion mechanism.

Smith (1994b) suggests that a feature tracking mechanism exists side by side with the lower level first- and second-order channels. He presented subjects with a contrast defined version of the 3*f* plus 4*f* stimulus employed by Hammett *et al* (1993). The stimulus was constructed by modulating the contrast of static binary noise rather than, as in the case of Hammett *et al*, modulating luminance. This stimulus is opaque to standard motion energy analysis (Chubb & Sperling, 1988) so any motion elicited by the stimulus must indicate the operation of some non-Fourier mechanism or mechanisms. Smith found that, with no ISI, reversed motion was perceived. This finding is fully consonant with the notion of a second-order channel in which the signal is demodulated and then subject to motion energy analysis. However subjects perceived motion in the forward direction when a short ISI was incorporated into the stimulus. Clearly this must reflect the operation of some process other than simple demodulation and subsequent standard motion energy analysis. Smith takes his findings to be evidence for a tripartite model of human motion perception in which some feature tracking mechanism exists in parallel to the first- and second-order mechanisms.

Lu & Sperling (1995b) also provide evidence that they interpret as confirming the existence of this tripartite architecture. Their basic strategy is to add a "pedestal" to sequences first- and second-order motion sequences where the pedestal is a static pattern with the same spatial frequency and configuration (ie. luminance or contrast modulation) as the motion sequence to which it is added. Tracking the features of this pattern simply gives an oscillation, whereas some motion energy mechanism applied to the luminance modulated version of the stimulus (or the demodulated contrast modulated stimulus) will signal motion

in the direction of translation. Lu & Sperling believe that the addition of the pedestal is therefore a potent technique by which to isolate motion systems that do not rely on feature tracking. Lu & Sperling show that motion is readily perceived in both luminance and contrast modulated pedestal stimuli. This is taken to demonstrate the existence of separate first- and second-order motion systems which are both ultimately based upon some motion-energy type motion detection process. Demodulation is incorporated into the second-order channel so that the motion is visible to standard motion energy analysis.

Lu & Sperling (1995b, 1995c) demonstrate the existence of a feature tracking channel by showing that the perception of motion elicited by some types of stimuli are severely disrupted by the addition of a pedestal. Most notably these include interocular displays. In these the motion information available to any single eye will not preferentially signal motion in any single direction. However, when information from both eyes is combined then motion should be visible. Lu & Sperling propose that the first- and second-order systems are primarily monocular and that interocular motion displays isolate the feature tracking mechanism. The conclusion that low level motion systems are primary monocular rests upon their finding that interocular motion is disrupted by the addition of a pedestal.

The issue of whether low level motion mechanisms extract interocular motion has been a point of debate for some years (Georgeson & Shackleton, 1989; Carney & Shadlen, 1992; Georgeson & Shackleton, 1992, Carney & Shadlen, 1993; Carney, 1995) however, with respect to the specific issue of whether interocular motion is disrupted by the addition of a pedestal, Carney (1997) produces convincing evidence to show that this is not necessarily the case. Carney presented two counterphasing sinusoidal gratings to his subjects (one to each eye). The gratings were in quadrature phase with one another so that by summing the two together a translating sinusoidal grating is produced. Carney showed that the motion in this stimulus could readily be perceived even with the addition of pedestals with contrasts greater than those of the motion stimuli to which they were added. The difference in findings between Carney and those of Lu & Sperling (1995b) is accounted for by appealing to the comparatively short stimulus presentation times employed in the latter study. Carney also showed that the direction of motion in interocular stimuli can be determined at temporal frequencies of up to at least 32 Hz. This observation is at odds with the assertion by Lu &

Sperling that the response if the feature tracking mechanism is low pass with respect to temporal frequency and has a cut-off frequency of approximately 3 Hz.

If feature tracking is indeed part of human motion perception then the mechanisms underlying it are little understood. A number of very basic questions remain to be answered. For example, what are the features that are selected, how are they extracted and how are they tracked? At present the most complete psychophysically based description of human feature tracking is provided by Lu & Sperling (1995b, 1995c), but it should be noted that the characterisation that they provide of the feature tracking mechanisms appears, on the basis of other evidence (Carney, 1997), somewhat dubious. Lu & Sperling characterise feature tracking as both an active and a passive process. They put forwards the theory that feature tracking is accomplished by mapping the features onto some saliency map which is then subject to standard motion energy analysis. They also note that some of their evidence does not totally concur with this theory. However, in spite of the generally unspecified nature of the putative feature tracking system, the notion of feature tracking as a general mechanism involved in human motion perception has become widely accepted.

1.4. Summary.

Although there are a number of low level intensity based models of human motion processing, these are generally perceived to be equivalent and to operate by recovering orientations of Fourier motion energy through the frequency domain origin. However, there exists a class of stimuli, generally termed second-order, that elicit a perception of motion which cannot be explained in these terms. It was initially suggested that a non-linearity, early in visual processing, introduced distortion products at the frequencies of the second-order motion into the processing stream. In this scenario, the perception of second-order motion is simply an accidental by-product created by a deviation from ideal linearity in the visual system. In the light of psychophysical evidence, this hypothesis cannot fully account for the perception of second-order motion. This has led to the proposal of a second-order processing stream that acts in parallel to the traditional first-order linear channel. In this channel the stimuli are subjected to some gross non-linearity, such as rectification or squaring which demodulates the second-order signal. Attempts to provide definite evidence for or against the two channel architecture have so far been inconclusive. The issue of whether a second-order channel really does exist is still a matter of debate.

The hypothesis that motion is detected through the operation of spatio-temporal energy detectors has come to dominate the field of motion perception. The theory has motivated a vast amount of psychophysical and physiological research. The whole issue of how we perceive second-order motion has a critical bearing on this core hypothesis. The study of the perception of second-order motion therefore has a direct bearing on the question of how we perceive motion. In this thesis the notion of a specific non-linear channel for second-order motion and the wider issue of non-linearities in second-order motion perception are examined. This is achieved through the techniques of psychophysical investigation and computational modelling. The results from the studies detailed in this thesis cast some doubt over the notion of a second-order channel and raise the possibility that the motion from Fourier components approach, as a conceptual framework for motion perception, may be in need of some revision.

Chapter Two

General Methods.

2. General Methods.

There is considerable variance in the experimental methodology across the experiments described in this thesis. Many of the specific methodological details are therefore described in the relevant chapters.

2.1. Equipment.

In the majority of the experiments, stimuli were displayed on a non-interlaced monochrome Manitron monitor, driven by a Matrox IM-1280 graphics card that was controlled by an IBM compatible micro-computer. The Manitron monitor was equipped with P33 phosphor. The displayable area of the screen had a width of 36.7 cm and a height of 29.3 cm. With the setup described above, the size of the screen contained 1280 by 1024 picture elements (pixels). The frame rate was 59.5 Hz. The graphics card had 8 bit output giving a total of 256 possible grey levels.

Two further experimental setups were utilised. In one, the Manitron monitor was driven by a Cambridge Research systems VSG 2/3 board controlled by an IBM micro-computer. This system was capable of 14 bit output giving a total of 16,384 grey levels. The frame rate was 60 Hz and the display area of the screen was 36.2 cm wide (656 pixels) and 25.7 cm high (908 pixels). In the other experimental setup, stimuli were produced and displayed on a Sun SparcStation LX. The display area of the screen had a width of 34.7 cm (1152 pixels) and a height of 27.0 cm (896 pixels). The frame rate was 60 Hz and the system was capable of displaying 256 grey levels.

2.2. Stimulus Generation.

Two general methods were employed for stimulus generation. In one, ramps were drawn in video memory and were then displayed on the screen after being transformed by an output look-up-table (LUT). By cycling different images through the LUT, a one dimensional translating image can be produced. This is best illustrated by taking a concrete example. To display a vertically oriented horizontally translating sine wave grating in an area of the screen that has a width of 256 pixels and a height of 128 pixels. First, a ramp running from

left to right (or vice versa) is created in graphics memory. The ramp has a width of 256 and a height of 128. The values in the ramp run from 0 to 255 in steps of 1. The ramp is then used to access the values in an output LUT. Any part of the ramp with a value *n* will access the *n*th value of the LUT. Therefore the cross section of the LUT, for example a sine wave, will be repeated down the screen to produce a vertically oriented sine wave grating. By feeding a space-time image onto the graphics card, and then by loading each frame of the space time image into the LUT, a translating sine wave grating can be produced. An advantage of this method is that the images utilise little graphics memory. Therefore long sequences (potentially hundreds of frames) can be stored on the graphics card and cycled through. The disadvantage is that this procedure cannot readily be used to produce images that vary in two spatial dimensions.

In the second method used to generate images, two dimensional frames were generated and stored in buffers that had been set up on the graphics card. These images were then cycled through to produce apparent motion sequences. With this technique, each complete two dimensional frame must be stored on the graphics card. The method is therefore far more memory intensive than the ramp plus LUT arrangement. The amount of memory used may be reduced by scaling up the image so that, for example, a 320 by 256 array on the graphics card may be scaled by a factor of 4 to produce a full screen 1280 by 1024 image. If this is the case then each "stimulus pixel" is comprised of 4 by 4 physical pixels. By increasing the scale factor, a larger number of frames may be stored on the graphics card. This method has the advantage that images that vary in two spatial dimensions may readily be produced.

2.3. Gamma correction.

In the following description the term "grey level value" is used to refer to the actual value present on the graphics card. For example, on the Matrox graphics board each image is represented by an array of 8 bit values in graphics memory, each of those 8 bit values may be termed a grey level value. Each grey level value produces a certain output voltage and each output voltage produces a certain luminance on the display. There are therefore two potential sources of non-linearity: firstly, the transformation of the grey level value into an output voltage; and secondly, the relationship between output voltage and luminance. Gamma correction was accomplished by measuring the display luminance produced by a

range of grey level values and fitting a function to the data so that the grey level value needed to produce a certain luminance could be calculated. In this way, both sources of non-linearity may be accounted for by a single operation. A number of devices were used to measure the luminance of the screens, these are detailed in the relevant chapters. For the Manitron/Matrox system and the Manitron/VSG system, a function of the form

$$p = k.I^{\gamma} \qquad 2.1$$

was fitted to the data (p is the grey level value, k is some constant and γ is a measure of the non-linearity). Taking the log of both sides of the equation gives the following:

$$\log(p) = \log(k) + \gamma.\log(I) \qquad 2.2$$

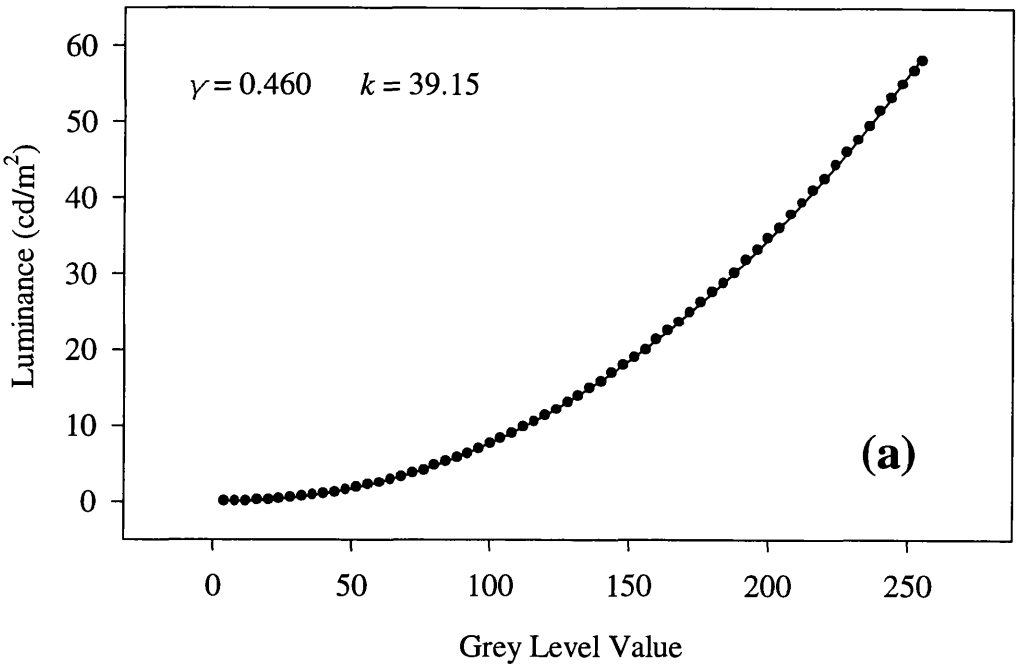
which means that the logged data can be fitted to the function by simple linear regression. An example of gamma correction for the Manitron/Matrox system is shown in Figure 2.1a. Luminance was measured using a Graseby (UDT) model 265 luminance probe. As one can see, equation 2.1 provides a good description of the data.

Having derived values for k and γ one can work out the grey level needed to produce any particular luminance value. In the majority of experiments the gamma correction process was applied during the calculation of the grey level values of the images stored on the graphics cards. For example, if one wished to produce a translating sine wave grating using the ramp plus LUT method described above, then the values in the space-time image would be calculated such that, when the ramp accessed the LUT, the luminance cross section of the image on the screen would be sinusoidal. The space-time image stored on the graphics card would contain a distorted version of a sine wave grating, the distortion would be dependent upon the gamma correction function.

For the Sun SparcStation LX, equation 2.1 did not provide an adequate fit for the data. In this case a cubic equation of the form

$$I = b_0 + b_1.p + b_2.p^2 + b_3.p^3 \qquad 2.3$$

was fitted (where I is luminance, p is the grey level value and b_0, b_1, b_2 and b_3 are constants). As one can see from Figure 2.1b, this equation provides a reasonable fit for the data. The



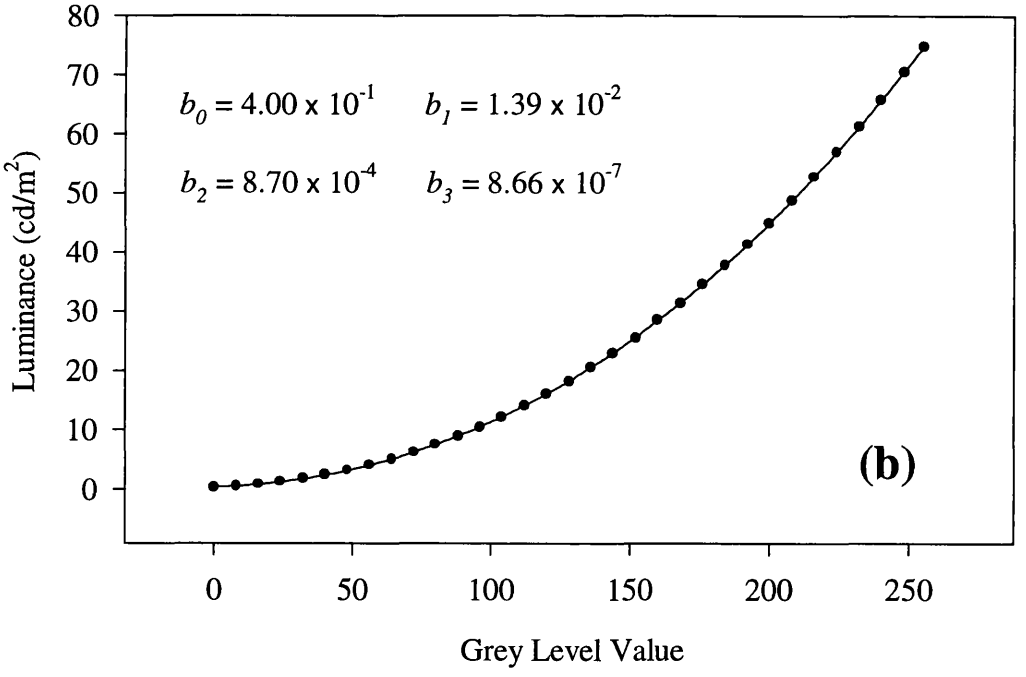


Figure 2.1: (a) An example of gamma correction for the Manitron/Matrox system, fitting equation 2.1 to luminance data. Values for k and γ determined by linear regression are shown on the graph. (b) An example of gamma correction for Sun SparcStation LX, fitting equation 2.3 to luminance data. Values for b_0, b_1, b_2 and b_3 are shown on the graph.

Grey Level Value	Calculated Luminance Value (cd/m^2)
130	18.817101
131	19.102400
132	19.390100
133	19.680201
134	19.972799
135	20.267799
136	20.565300
137	20.865200
138	21.167601
139	21.472401
140	21.779699

Table 2.1: List of grey level values and associated luminances. The luminance values were calculated by fitting equation 2.3 to the data in Figure 2.1b and then predicting the luminance from the fitted equation.

only problem with this formulation is that it is not easy to find what particular value of p is needed to produce a particular value of I. To evade this problem, the following procedure was used. For all the values of p from 0 to 255 a list of the respective luminances was calculated. A portion of this list is shown in Table 2.1. To take a concrete example, suppose a luminance of 20.4 cd/m^2 was required. The nearest grey levels on either side would be found by applying some search procedure to the list. A grey level of 135 gives a luminance value of 20.2678 cd/m^2 whilst a grey level of 136 gives a luminance value of 20.5653 cd/m^2. One simple way of finding the necessary grey level is simply to use the one that will produce the closest luminance to that which is required. In the example given above this means that a grey level value of 135 would be used. However in the experiments described in this thesis, a more complex procedure was employed. This is described in the following paragraph.

The choice of grey level value would be determined probabilistically in the following manner. If I_{want} is the required luminance value and I_{low} is the luminance produced by the nearest lower grey level and I_{high} is the luminance produced by the nearest higher grey level then some value n exists such that:

$$I_{want} = n.I_{high} + (1 - n).I_{low} \qquad \text{(where } 0 \leq n \leq 1\text{).} \qquad 2.4$$

The value of n is given by the equation

$$n = \frac{I_{want} - I_{low}}{I_{high} - I_{low}} \quad . \qquad 2.5$$

If the probability of choosing the grey value associated with I_{high} is n and the probability of choosing the grey value associated with I_{low} is 1 - n, then as the number of choices increases, the mean luminance of those choices will tend towards I_{want}.

2.4. Psychophysical techniques.

The majority of the experiments utilised the Method of Constants. In some experiments, the levels and number of trials per level were predetermined; in others, the values of the stimulus variables with which the subjects were tested were chosen adaptively from a set of possible values. In all but one instance the adaptive procedure was similar to the Adaptive Probit Estimation (APE) procedure described by Watt & Andrews (1981). In the following description an experimental run consists of a number of trials where each trial is the presentation of a single stimulus (ie. with a method of single stimulus task) or pair of stimuli (ie. with a spatial or temporal two alternative forced choice task). The value of the experimental variable is chosen on an estimate of the mean and standard deviation of a cumulative Gaussian fitted to the data collected from the previous 32 trials in the run. Of course the procedure is different when less than 32 trials have been completed within the run. If the number of trials is between 16 and 32 then a psychometric function is fitted to all of the data that has so far been gathered. When less than 16 trials have been completed then the value of the experimental variable is derived from an estimate of the mean and standard deviation of the function provided by the experimenter at the start of the run. In the original Watt & Andrews (1981) formulation, a regression would be performed at the end of a block containing a fixed number of trials and the values presented in the next block would be dependent upon the estimated mean and standard deviation. The use of the sliding window in the current version of APE is described by Treutwein (1995). The procedure used for the

curve fitting was probit analysis as described by Finney (1971) except that the regressions were unweighted. After a specific number of trials (typically 64), the run is completed and standard probit analysis (with weighted regressions) is applied to the complete data set. This procedure allows both the mean and the standard deviation of the psychometric function to be estimated.

Chapter Three

First-order motion from contrast modulated noise?

3. First-order motion from contrast modulated noise?

The class of microbalanced motion stimuli is thought to contain no systematic directional biases in motion energy (Chubb & Sperling, 1988). The fact that we can see motion in such stimuli implies that models of human motion perception based on Fourier decomposition need to be revised. The validity of one widely studied class of microbalanced stimuli, contrast modulated noise, has recently been questioned (Smith & Ledgeway, 1997a). It has been proposed that stochastic local biases in the noise carrier give rise to luminance artifacts detectable by a Fourier energy mechanism. All of the second-order stimuli employed in the following chapters are contrast modulations of noise. The issue of whether the motion in such stimuli can be extracted by first-order mechanisms is therefore critically important for any conclusions that can be drawn from the results of the experiments described in this thesis. In this chapter I show that the response of a motion energy system to contrast modulated noise shows no directional bias over a number of carrier configurations. This implies that any consistent and coherent motion percept elicited by contrast modulated noise stimuli must reflect the operation of some non-Fourier system.

3.1. Introduction.

Microbalanced stimuli were initially described by Chubb & Sperling (1988). The most common examples of stimuli of this type are contrast modulations of binary noise (see Figures 3.1 & 3.2). These are often termed *second-order*, as they involve motion defined by second-order statistics of the image (Cavanagh & Mather, 1989). A stimulus may be termed *microbalanced* if the result of windowing that stimulus with any space-time separable function is *drift-balanced.* Chubb & Sperling (1988) define a stimulus as drift-balanced if its expected power within the frequency domain is symmetric with respect to temporal frequency. It is for this reason that the motion elicited by these stimuli has been termed *non-Fourier.*

A number of influential models (Adelson & Bergen, 1985; Watson & Ahumada, 1983; van Santen & Sperling, 1984) are seen to extract motion by locally detecting the orientations of Fourier energy present in the signal (Adelson & Bergen, 1985, 1986). These models detect *first-order* motion as they operate on modulations of luminance, a first-order characteristic.

The existence of non-Fourier motion implies that either the principle of deriving motion from Fourier components is incorrect, or some additional non-Fourier mechanism exists. Because microbalanced stimuli have such a strong bearing on models of human motion processing, they have elicited a great deal of research. The most convenient stimulus of this type has proved to be contrast modulated noise.

Smith & Ledgeway (1997a) make the point that there may well be biases in the ratio of light to dark pixels over a given area. They make the seemingly reasonable assumption that in areas where light or dark pixels do predominate there will be some degree of luminance modulation (see Figure 3.1). This latter point can be exemplified by taking the idea to its extreme. Within a given area there is obviously some finite chance that all of the noise pixels will be either light or dark. If this does occur then, within that area, the stimulus will effectively be luminance defined rather than contrast defined. If this is the case then standard motion energy analysis will serve to extract motion in this region of the stimulus.

These stimuli may therefore be seen as consisting of areas which can vary on a continuum between regions where there are equal numbers of light and dark pixels, and regions in which pixels of one type dominate. At one end of the spectrum we have the contrast defined ideal where no luminance defined motion is present. At the other end of the spectrum we have the likelihood of contamination by luminance artifacts. In a given spatio-temporal window the probability of significant deviation from equal numbers of light and dark pixels is increased if the number of pixels within that window is decreased. Therefore by increasing the size of the pixels or decreasing the dimensionality of the noise (ie. from dynamic to static noise) we should increase the probability of contamination by luminance artifacts. If the analysis provided by Smith & Ledgeway (1997a) is correct, then contrast modulated static noise may be particularly unsuitable for use in experiments on second-order motion. This would also mean that a portion of the literature on the subject needs to be re-evaluated. In this chapter the theory that luminance artifacts can bias the detection of motion by a Fourier energy system is examined.

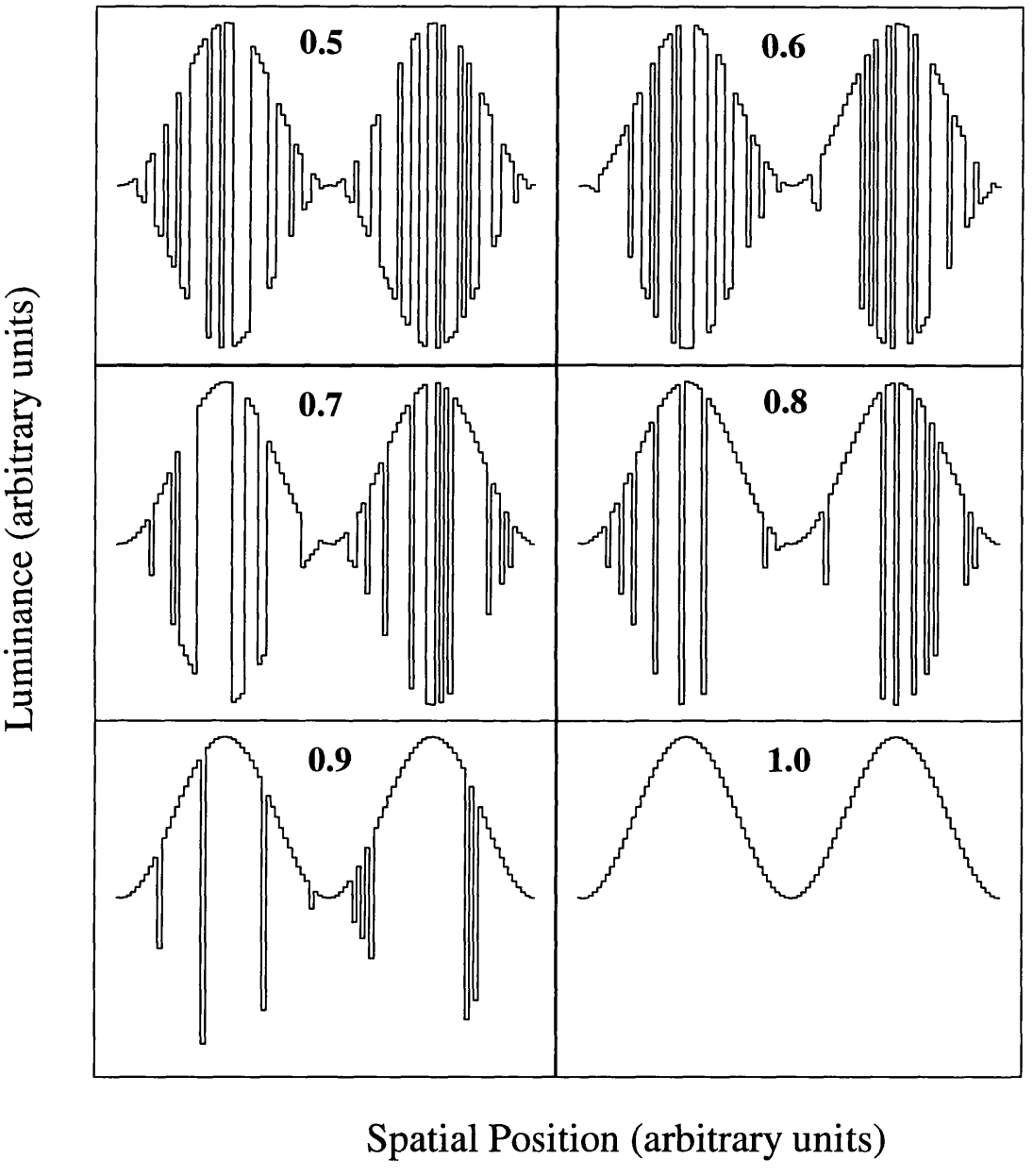


Figure 3.1: Spatial cross sections from some contrast modulated noise stimuli. The figure at the top centre of each graph shows the probability of a bright pixel occurring in the underlying noise carrier. As the probability increases the stimulus becomes increasingly luminance-defined rather than contrast-defined.

3.2. Method and Results.

A standard energy model was applied to instantiations of contrast modulated noise stimuli and the response of the model to these stimuli was examined. The energy model utilised two quadrature pairs of space-time oriented Gabor filters. The two pairs were tuned to the same absolute spatial and temporal frequencies but responded maximally to motion in opposite directions to one another. The outputs were squared and added within each quadrature pair to gain a directionally sensitive phase invariant measure of local motion energy. By taking the difference between the phase invariant outputs, a local measure of opponent motion energy was extracted (see Adelson & Bergen, 1985).

The model was applied to stimuli consisting of contrast modulated binary random noise. The input images were 480 pixels square. So as to maximise any potential detection of motion energy, the spatial and temporal frequencies comprising the modulant in the stimulus were identical to those of the sine waves making up the Gabor filters of the model. For both spatial and temporal frequencies this was set to 10 cycles per image, giving a wavelength of 48 pixels. Each output image consisted of measures of opponent energy at 278^2 locations. The model was tested over a range of noise element sizes (1, 2, 3, 4, 6, 8 and 12 pixels) with both static and dynamic noise carriers. The noise size represents the spatial width of the noise. The temporal extent of the dynamic noise was 1 pixel.

Examples of some input images, and the output of the model in response to them, are shown in Figure 3.2. It is clear that these stimuli do produce a response in the motion energy model. It is also clear that both forwards *and* reversed motion are detected in contrast modulations of both static and dynamic noise. Figure 3.3 shows a section of input image containing areas that elicit forwards motion and areas that elicit reversed motion. We can see that the reversed motion occurs in regions where one finds reversals in the sign of the carrier. Figure 3.3 shows how a space-time oriented filter applied to these areas of the image will preferentially signal reversed motion. It is not clear from the individual raw output images whether forwards or reversed motion might be dominant.

The model was applied to 100 instantiations of each input image. For every instantiation the noise carrier was freshly generated. For each output field the sum of the responses, and the

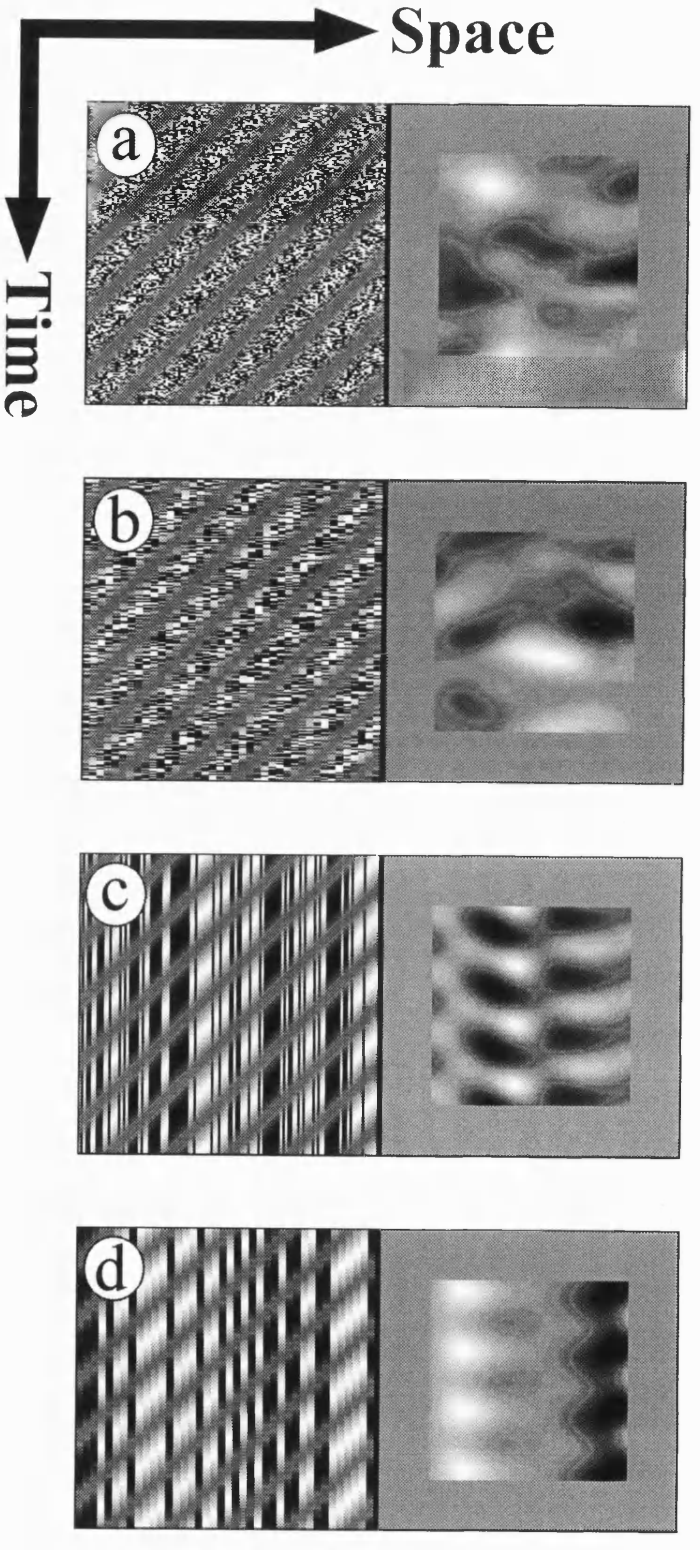


Figure 3.2: Space-time plots showing the output of an energy model applied to contrast modulated noise stimuli with (a) dynamic noise carrier, noise element size of 2 pixels; (b) dynamic carrier, element size of 6 pixels; (c) static carrier, element size of 2 pixels; and (d) static carrier, element size of 6 pixels. Input is shown on the left, output on the right. In the output images, areas lighter than the grey border indicate leftwards motion, areas darker than the border indicate rightwards/reversed motion.

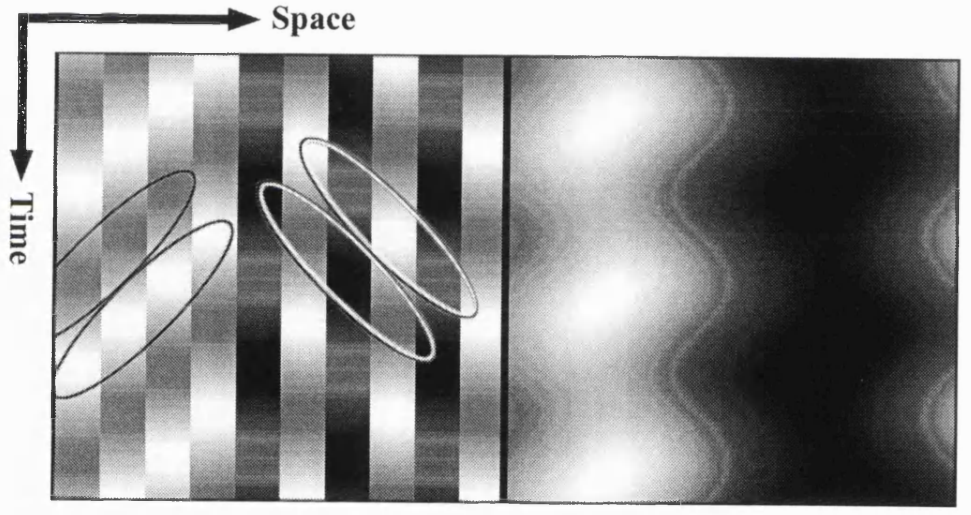


Figure 3.3: Space-time plots showing a section of input image and its associated output. The left hand side of the input image signals leftwards motion and the right hand side of the image signals rightwards motion. The oriented oval fields indicate how space-time oriented receptive fields might be applied to the image. The output response increases as the difference in mean luminance between the adjacent lobes increases. For maximum response, the fields in the right hand side of the input image must be oriented such that they signal motion in a direction opposite to that of modulant translation.

sum of the squares of the responses, were extracted. This allows the mean and standard deviation collapsed over both area and instantiations to be calculated. In other words, one can calculate the mean and standard deviation of responses gathered from 100 output fields, each of which contain 278^2 responses. Each mean and standard deviation is therefore drawn from approximately 7¾ million measures of opponent energy. These data are plotted in Figure 3.4. It is clear that there is no apparent overall directional bias in the measures of opponent energy. Increases in noise element size lead to increases in the standard deviation of the response. The relationship appears quasilinear. The standard deviations obtained in response to static noise carriers are greater that those obtained in response to dynamic noise carriers by a factor of 10.

3.3. Discussion.

The data show that both increasing noise size, and the use of static rather than dynamic noise carriers, increase the likelihood of local luminance imbalances occurring. What the data also show is that the probability of these imbalances signalling motion is directionally balanced. In other words, for any single instantiation of a contrast modulated binary noise stimulus, there can be no *a priori* expectation of motion direction in terms of local measures of motion energy. If motion in these stimuli is consistently perceived in a certain direction then this must be due to the operation of some process other than motion energy extraction.

Smith & Ledgeway (1997a) examined orientation and direction discrimination thresholds for contrast modulations of binary noise. With dynamic noise carriers, direction discrimination was considerably worse than orientation discrimination. When static carriers were utilised, this difference was far less pronounced. Smith & Ledgeway suggested that the raised threshold for direction discrimination characterised the operation of a second-order system. They argued that this characteristic was not elicited by modulations of static noise because the motion of the modulant in these stimuli could be detected by a first-order system. The results described in this chapter have shown that this assertion cannot be sustained. Unless one is to suppose that separate second-order mechanisms exist for the detection of modulations of static and dynamic carriers, the pattern of results must be explained without recourse to an extra mechanism.

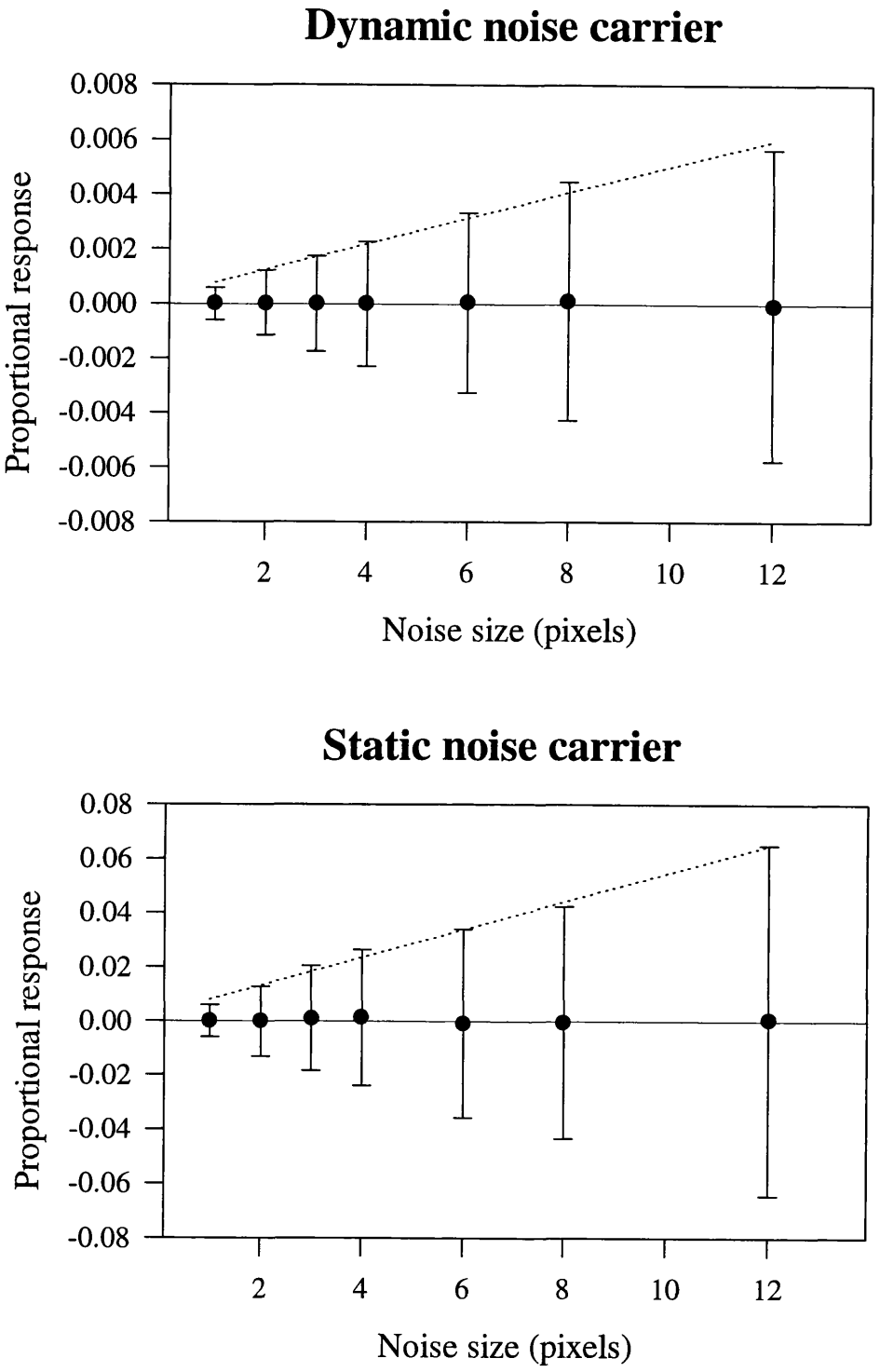


Figure 3.4: Graphs showing means and associated standard deviations of responses of an energy model to contrast modulated noise stimuli (see text for details). The *y* axes are scaled by the response of the energy model to a translating sine wave grating of the same spatial and temporal frequency as the modulant in the contrast modulated stimuli. This sine wave had the same maximum and minimum values as the contrast modulated stimuli. Note the change of scale between the two graphs. The dotted lines are regression lines for the standard deviations; correlation coefficients are as follows: dynamic noise, $r^2 = 0.989$; static noise, $r^2 = 0.995$.

One possible explanation may lie in the differences between the first-order components of the two types of motion pattern. Figure 3.5 shows "mean" Fourier spectra for contrast modulations of static and dynamic noise. These were created by averaging the Fourier spectra of 100 instantiations of each stimulus type. The energy components that carry motion direction information are those that do not lie on the lines in the spectra representing a temporal frequency of zero or a spatial frequency of zero. In the case of the dynamic noise carrier it is clear that a far greater proportion of the energy present in the image carries motion direction information than in the case of the static carrier (in fact 98.8% compared to 49.7%). Modulations of dynamic noise therefore contain a higher proportion of what might be termed "motion direction noise" than modulations of static carriers. It seems reasonable to suggest that increased motion direction noise may cause increased thresholds in a direction discrimination task but may not necessarily affect performance on an orientation discrimination task. The differences in performance between contrast modulations of static and dynamic noise may therefore reflect an interaction between the nature of the tasks and the nature of the stimuli rather than the operation of two separate systems.

3.4. Conclusion.

The simulations support the analysis offered by Chubb & Sperling (1988). An analysis of results detailed by Smith & Ledgeway (1997a) is offered without resorting to a dual mechanism approach. If this analysis is correct, then a wider implication of the results from this chapter is that differences in performance elicited by different stimuli do not necessarily indicate the operation of two or more separate mechanisms. On a practical level, the findings show that contrast modulated noise remains a potent tool with which to examine the central issue of whether human motion processing proceeds by the extraction of motion from Fourier components, and also for examining the characteristics of any postulated non-Fourier mechanisms. Any coherent percept of motion elicited by contrast modulated noise stimuli must reflect the operation of a non-Fourier system.

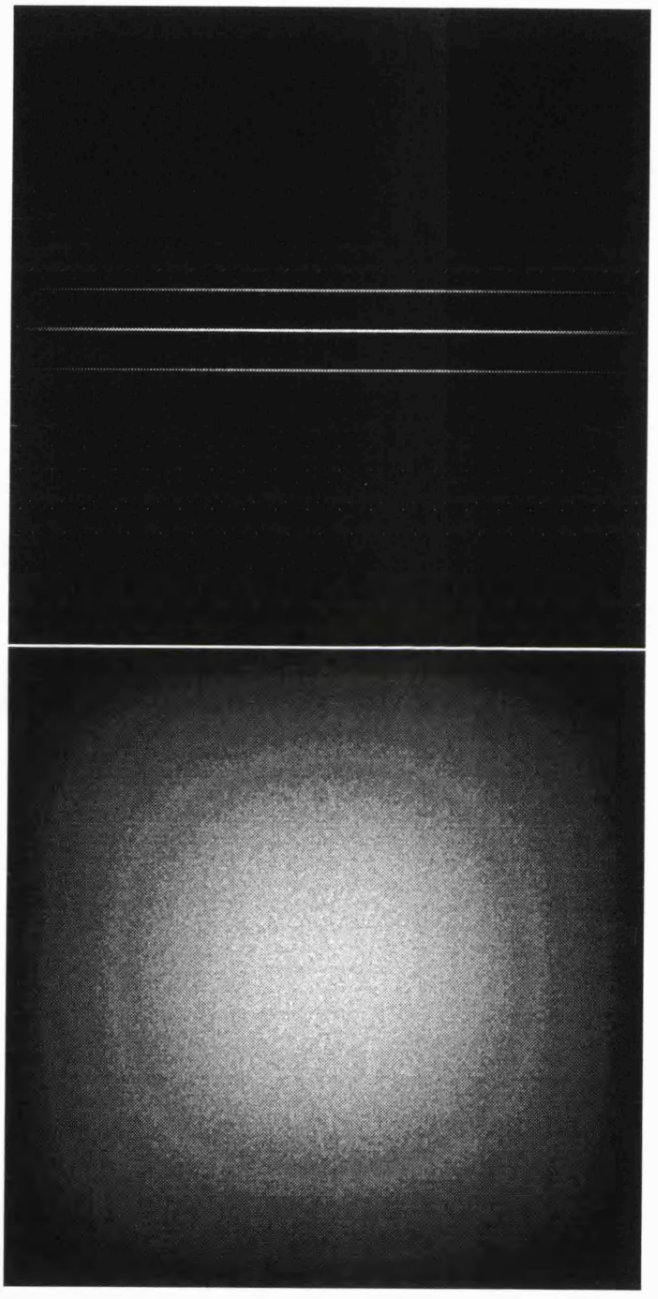


Figure 3.5: Averaged Fourier spectra of 100 instantiations of a contrast modulated static noise stimulus (top) and 100 instantiations of a contrast modulated dynamic noise stimulus (bottom). The spatial and temporal frequencies of the modulants were set to 16 cycles per image. The noise width was set to an eighth of a cycle.

Chapter Four

Speed discrimination thresholds for first- and second-order bars and edges.

4. Speed discrimination thresholds for first- and second-order bars and edges.

Speed discrimination thresholds were measured for first- and second-order bars and edges as functions of stimulus speed, width and magnitude. In terms of the pattern of results over speed and width, there was no evidence for any systematic differences between first- and second-order stimuli. In general, thresholds were higher for second-order stimuli than for first-order stimuli. This difference was more pronounced between the first- and second-order bars than between the first- and second-order edges. It is argued that the data can be accounted for through the operation of a single mechanism for the extraction of first- and second-order velocity.

4.1. Introduction.

Within the motion from Fourier components framework, frequency space is seen as being sampled by a mosaic of spatio-temporal motion energy detectors. Because the output of a motion energy detector is dependent upon the contrast, spatial frequency and temporal frequency of the input, it is not possible to extract velocity from the output of any single detector. There are a number of schemes for extracting velocity from energy models (Adelson & Bergen, 1985; Heeger, 1987; Grzywacz & Yuille, 1990), all of which involve some comparison of the output between two or more motion energy detectors. Velocity extraction is therefore accomplished by a two stage mechanism. Firstly, the detection of motion by directionally selective units and, secondly, some integration process applied across the outputs of a number of units.

Speed discrimination thresholds (SDTs) give us a measure of the resolution of the velocity extraction system. Given the mechanism outlined in the paragraph above, this measure should be dependent upon both stages of the mechanism. Within the two channel model, second-order motion is extracted by demodulation followed by motion energy detection. The motion energy detectors in the second-order channel are seen as separate from those that operate in the first-order channel. If the spatio-temporal mosaic of motion energy detectors in the first- and second-order systems are preferentially tuned to different spatial and/or temporal frequencies, this should be reflected in the patterns of speed discrimination thresholds over spatial and temporal frequency. Unless we assume that frequency space is

sampled in the same way in both channels, it would seem reasonable to assume that SDTs should follow dissimilar patterns of response to changes in the temporal and spatial parameters of first- and second-order stimuli.

There has been a considerable amount of research into speed discrimination thresholds in first-order motion perception. The basic findings are that, in the fovea and over a range of velocities from $2°s^{-1}$ to $64°s^{-1}$, SDTs range between 0.04 and 0.1 (all SDTs will henceforth be given as Weber fractions) and that SDTs plotted against base velocities follow a U-shaped relationship. This has been shown with random dot patterns (De Bruyn & Orban, 1988) and luminance bars (Orban, De Wolf & Maes, 1984; Orban, Calenbergh, De Bruyn & Maes, 1985) over a wide range of target velocities. With increasing eccentricity the lower end of the U-shaped function is shifted towards higher velocities and the velocity at which optimum SDTs occur is increased. However, the upper limit appears to remain constant (De Bruyn & Orban, 1988). A number of studies have looked at velocity discrimination in sine wave gratings (Thompson, 1983; McKee, Silverman & Nakayama, 1986; Smith, 1987; Panish, 1988; Smith & Edgar, 1991) and Panish (1988) additionally examined SDTs in Gaussian luminance bars. Although the velocity ranges employed in these studies are smaller than those used by Orban (see above), the basic findings in terms of levels of SDTs and relationship between SDTs and base velocity are reproduced.

One prominent finding in the velocity discrimination literature is that SDTs are little affected by profound changes in stimulus parameters other than base velocity. For example, when looking at random dot kinematograms, Watamaniuk, Grzywacz & Yuille (1993) found that motion discrimination thresholds remained constant even over abrupt changes in dot density, a manipulation that appeared to change perceived speed. McKee *et al* (1986) showed that SDTs appear to be insensitive to fluctuations over a wide range of suprathreshold stimulus contrasts and Orban *et al* (1984), measuring velocity discrimination in bars of differing lengths, found little difference in SDTs for bars of length 7° and bars of length 1°. Snowden & Braddick (1991) looked at the effect of stimulus duration on SDTs in random dot patterns and found a strong effect, but only with very brief presentations (below 120msec). De Bruyn & Orban (1988) studied the effect of stimulus duration over a larger range of velocities ($1°s^{-1}$ to $256°s^{-1}$) and concluded that, for all velocities tested, optimal SDTs are reached with stimulus durations above 200msec. It would seem reasonable to conclude that SDTs in

general provide a durable and robust measure that is little affected by stimulus parameters other than velocity itself.

Turano & Pantle (1989) examined velocity discrimination thresholds in contrast modulated and sine wave gratings. They showed that, at contrasts and modulation depths 5 to 10 times above their respective detection thresholds, there appears to be little difference in SDTs between the two types of stimulus. SDTs for the contrast modulated gratings appeared to be slightly higher, although this effect was more pronounced at low velocities ($0.75°s^{-1}$ and $1.5°s^{-1}$) and seemed to disappear altogether in the intermediate velocity range ($3°s^{-1}$, $6°s^{-1}$ and $12°s^{-1}$). In terms of the shape of the curves produced, there was little difference between the two stimulus types and Turano & Pantle state that "velocity discrimination with the two types of stimuli appears equivalent" (Turano & Pantle, 1989, p218). On the basis of this observation they conclude that their results support a hypothesis put forward by Henning *et al* (1975) in which elements are arranged to be sensitive to both low frequency luminance modulations and to low frequency contrast modulations of a high frequency carrier.

Cropper (1994) studied velocity discrimination thresholds in luminance gratings and beats. The latter were created by the addition of two translating luminance gratings of different spatial frequencies. This was equivalent to a translating contrast reversing envelope modulating the contrast of a stationary sine wave carrier. The stimuli were presented in multiples of their contrast detection thresholds over a range of velocities from 0.25 to 4.0 degrees per second. At high contrasts there appeared to be no difference in performance between the two types of stimulus, a finding that agrees with those presented by Turano & Pantle (1989). However, at low stimulus contrasts, SDTs for beats were much higher than those for luminance modulations. Cropper notes that the profound decrease in speed discrimination thresholds with increases in beat contrast does not parallel results obtained in velocity discrimination studies of beats. He concludes that this may imply the speed and direction of motion in beats are extracted either by different mechanisms or are available at different stages within the same mechanism.

The use of stimuli constructed from modulating the contrast of binary noise has become widespread in the investigation of second-order motion perception. Such stimuli prove especially problematic for first-order motion extraction mechanisms (Chubb & Sperling,

1988). The motion evident in such microbalanced stimuli cannot be extracted through first-order motion processing. Some non-Fourier mechanism or process must be invoked to account for the perception of motion in these stimuli. Previous studies comparing SDTs in first- and second-order stimuli have employed contrast modulated sine wave gratings (Turano & Pantle, 1989) and beats (Cropper, 1994) as second-order stimuli. As these are not microbalanced, it is possible that the responses to the stimuli are based upon standard motion energy detection. In this chapter, investigations of second-order speed discrimination are extended to include microbalanced stimuli. Thresholds for two types of non-periodic first- and second-order stimuli are compared over stimulus velocity, stimulus width and stimulus amplitude.

4.2. Method.

4.2.1. Equipment.

Gamma corrected space-time images were constructed in PC RAM and passed to a Matrox Image-1280 graphics card. Data was loaded, frame by frame, from the space-time image into an output look-up table indexed by a ramp drawn in display memory. Each horizontal line in the space-time image (see Figures 4.1 and 4.2) represents one frame which is effectively constructed by spatially repeating that line across a region of the display device. The graphics card delivered 8 bits per pixel to give 256 grey levels. Images were displayed on a Manitron monitor. Slightly different versions of the same stimulus were presented simultaneously in two vertically aligned windows to the left and right of a fixation spot. The distance between the fixation spot and the centre of each window was 1.42° of visual angle. The windows themselves were square, each side measured 2.10°. In experiments 4.1 and 4.2 stimuli were 128 frames long whilst in experiments 4.3 to 4.6 the stimuli were 64 frames long. The frame rate was 59.5 Hz. There was a minimum of a 1 second gap between stimulus presentations. In each trial the screen around the stimulus was set to mean luminance (15.5 cd/m^2) and the full screen reverted to mean luminance between trials.

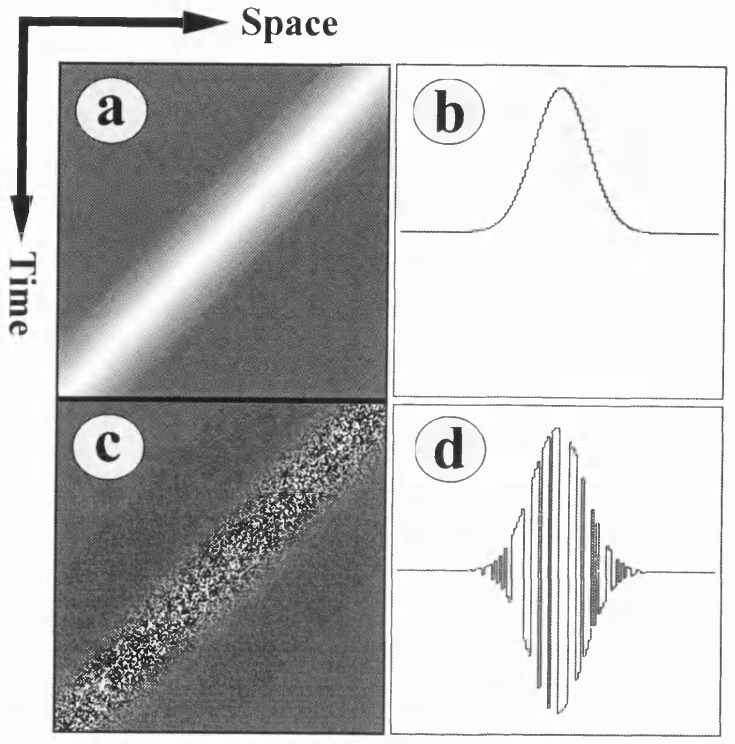


Figure 4.1: (a) Space-time plot of a first-order Gaussian bar. (b) A cross-section of the spatial luminance profile of a first-order Gaussian bar. (c) Space-time plot of a second-order Gaussian bar. (d) A cross-section of the spatial luminance profile of a second-order Gaussian bar.

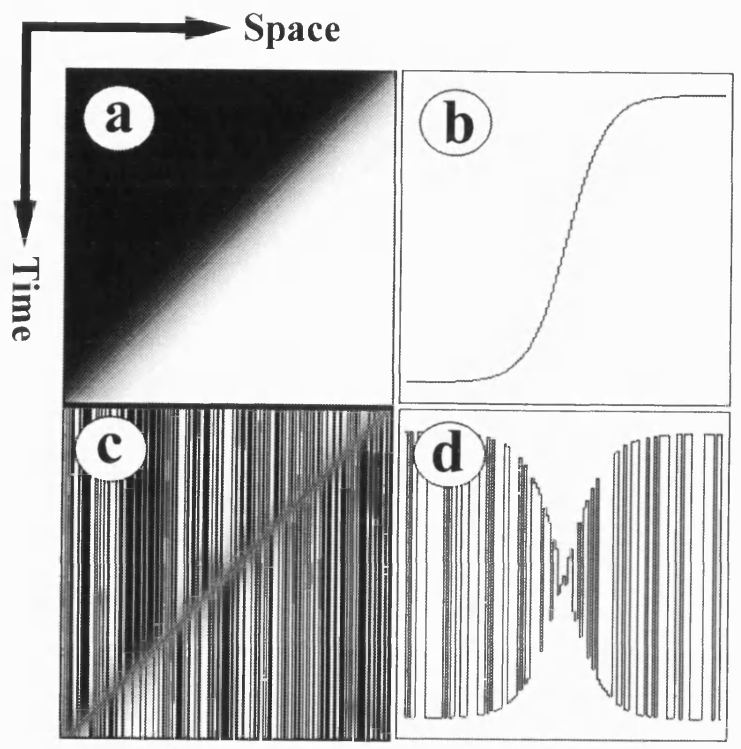


Figure 4.2: (a) Space-time plot of a first-order Gaussian edge. (b) A cross-section of the spatial luminance profile of a first-order Gaussian edge. (c) Space-time plot of a second-order Gaussian edge. (d) A cross-section of the spatial luminance profile of a second-order Gaussian edge.

4.2.2. Stimuli.

First-order bars: This stimulus consisted of a smoothly translating Gaussian bar. The minimum luminance in this image was set to the mean luminance of the display (15.5 cd/m^2), the maximum luminance in the image was set to 30.0 cd/m^2 giving a Michelson contrast of 0.32. A space time image of this stimulus is shown in Figure 4.1a and a spatial cross section of the luminance profile (in terms of deviation from mean luminance) is shown in Figure 4.1b. Measurements of bar width are given in terms of the spatial standard deviation of the Gaussian.

Second-order bars: A spatial Gaussian bar was used to modulate the contrast of binary 1D dynamic noise. The maximum luminance in the image was 30.0 cd/m^2, the minimum luminance was 1.0 cd/m^2. The contrast of the bar therefore varied from a maximum of 0.94 in the centre of the bar to zero at the edges. A space-time image and luminance profile of this stimulus are show in Figures 4.1c and 4.1d respectively. Measurements of bar width are given in terms of the spatial standard deviation of the Gaussian envelope. The noise had a width of 0.49 arc minutes.

First-order edges: This stimulus consisted of a smoothly translating edge which varied between 30.0 cd/m^2 and 1.0 cd/m^2, giving a Michelson contrast of 0.94. The luminance profile of the edge closely approximates the integral of a Gaussian. Measurements of edge width are given in terms of the spatial standard deviation of the underlying Gaussian. A space-time image and luminance profile of this stimulus are shown in Figures 4.2a and 4.2b respectively.

Second-order edges: A smoothly moving edge modulated the contrast of static 1D noise. The profile of the edge closely approximated the integral of a Gaussian. The edge was defined such that, in terms of contrast, it runs from a minimum of -0.94 to a maximum of +0.94. Although it is generally inappropriate to talk in terms of negative contrast, in this particular instance a negative contrast simply means that the polarity (in terms of deviation from display mean luminance) of the dark and light noise pattern is reversed and the contrast is set to its absolute value. The description is best conveyed by example; Figures 4.2c and

4.2d show a space-time image and luminance profile of the stimulus. The maximum luminance in the stimulus was 30.0 cd/m^2, the minimum luminance was 1.0 cd/m^2.

4.2.3. Spatio-temporal clipping.

A potential problem with speed discrimination is that judgements of relative speed can potentially be based upon a large number of factors other than velocity. McKee & Watamaniuk (1994) identify 4 factors that may covary with velocity, these being; stimulus duration, distance covered, contrast and temporal frequency. Whilst Weber fractions for distance discriminations are generally lower that those for speed discriminations (Burbeck, 1987) this may be controlled for by randomly varying the distance over which the stimulus moves. Differences in duration and temporal frequency are less discriminable than differences in velocity (Orban *et al*, 1984 and McKee *et al*, 1986). Therefore, at least at mid-range speeds where optimum discrimination is obtained, SDTs should be based on the percept of velocity.

In this study a number of random factors were built into the stimuli in order to reduce the likelihood that subjects could use stimulus duration or distance traversed as cues for speed discrimination. Two types of variation were used: spatial clipping and temporal clipping (see Figure 4.3). These ensure that, whatever the stimulus velocity, there will be some degree of randomness in the start position and spatial/temporal extent of the stimuli. Of course anything that affects stimulus duration for a non-periodic stimulus will also affect the stimulus duration. However, with the experiments described in this chapter, spatial clipping will preferentially apply to relatively slowly translating stimuli whilst temporal clipping will preferentially apply to more rapidly moving stimuli.

The way in which spatial and temporal clipping was applied differed for the two main stimulus types. For example, in Figure 4.3 the areas above b_1, below b_2, to the left of a_1 and to the right of a_2 would all be set to display mean luminance in the case of first- and second-order bars. In the case of the first- and second-order edge stimuli, if the edge is clipped then this simply determines the start or end positions of the movement of the edge. The edge remains on the screen for the entire stimulus duration but during some of that time it may be stationary. When a bar stimulus appears it will always be moving. After the clipped areas

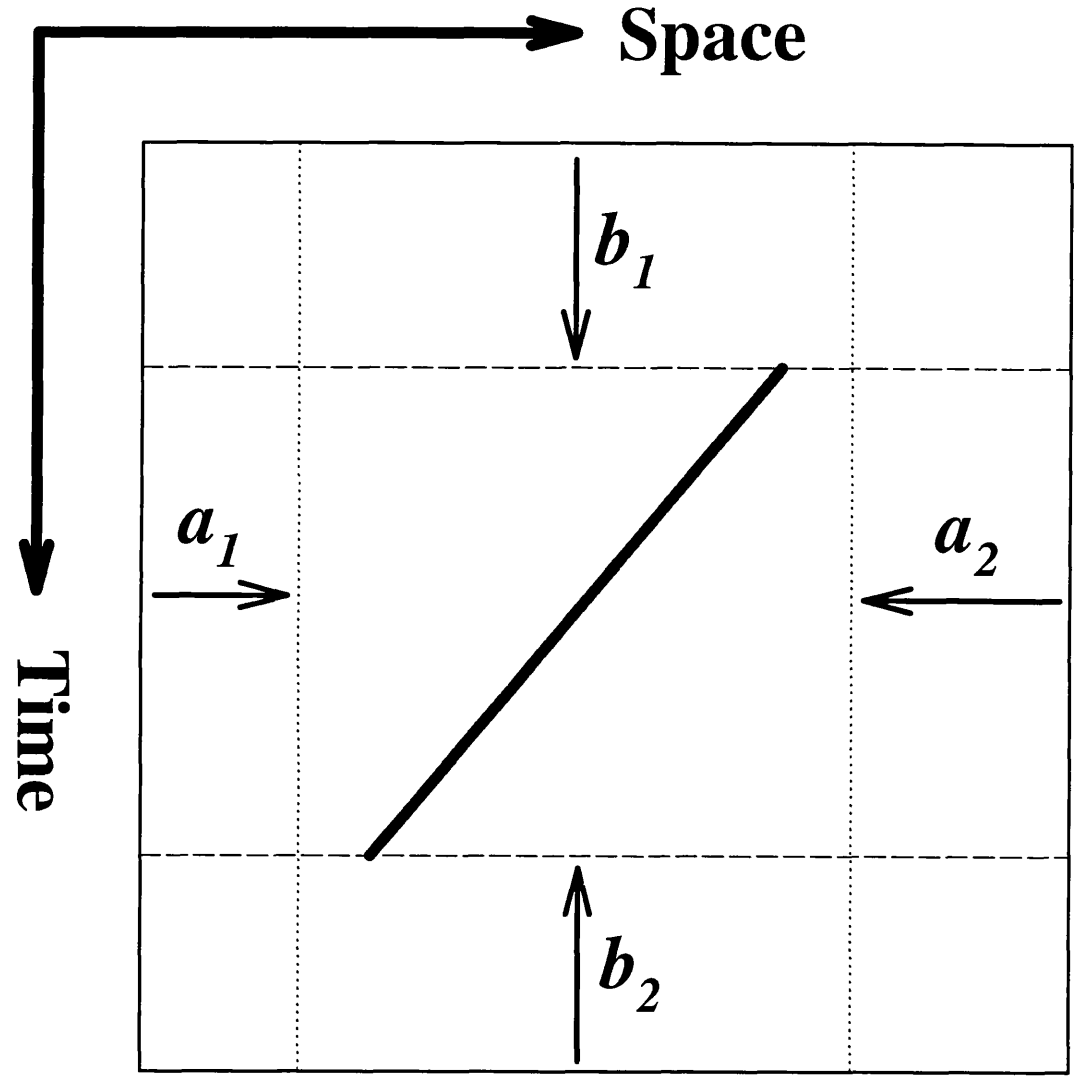


Figure 4.3: Parameters employed to clip the spatial and temporal extents of the stimuli. The values for the parameters are shown in Table 4.1.

have been determined the bar or edge is positioned so that it always goes through the centre of the remaining unclipped rectangle.

Spatio-temporal clipping was incorporated into the stimulus design so that if subjects did depend on either duration or distance, then the resultant thresholds should be higher than those commonly obtained in speed discrimination tasks. Table 4.1 shows the values over which the random values can range for the four experiments. The distribution of random values within the allowable ranges was flat. It should be noted that these parameters change the maximum and minimum translations and durations of the stimuli displayed in these experiments. For example in experiments 4.3 to 4.6, the duration of the stimuli was, on average, 32 frames (ie. ≈ 540 msec). This could vary between 42 frames and 22 frames, which gives a standard deviation of 6.20. Therefore if subjects based their decisions only on temporal duration, the very best performance that they could achieve (acting as perfect temporal discriminators) would be thresholds of approximately 0.2.

Experiment	*4.1 & 4.2*	*4.3 → 4.6*
Spatial Clipping (dva)	0.0 → 0.54	none
Temporal clipping (frames)	0 → 32	11 → 21
Number of Frames	128	64

Table 4.1: Ranges of values for parameters shown in Figure 4.3. Spatial clipping is given in terms of degrees of visual angle, temporal clipping is given in the number of frames.

4.2.4. Procedure.

The experiments described in this chapter compare velocity discrimination thresholds over first- and second-order stimuli as a function of either stimulus velocity, stimulus width or stimulus amplitude. Thresholds were determined using the Method of Single Stimulus (MOSS) combined with an Adaptive Method of Constants procedure (Watt & Andrews, 1981). Each experiment consisted of a number of levels, each level being the particular value of stimulus speed, width, or contrast/modulation depth that the subject was being tested upon. Each level contained a set of two randomly intermixed runs. The stimulus parameters were identical in both runs except that one of the runs contained first-order stimuli whilst the other contained second-order stimuli. Each run consisted of 64 trials, where a trial is a

single presentation of the stimulus. Prior to the presentation of each interleaved pair of runs, subjects were shown a number of standard stimuli and were given a short training set.

To take a concrete example, suppose that velocity discrimination in first- and second-order bars at $2°s^{-1}$ was being measured. Subjects were first presented with the standard stimuli which, in this case, were examples of first- and second-order bars moving at the base velocity of $2°s^{-1}$. Subjects were shown two of each of these (randomly shuffled together), no response was required. Subjects were then presented with a training set which contained 10 first- and 10 second-order stimuli that were moving either faster or slower than the standard stimuli. During the training set and the subsequent pair of interleaved runs, auditory feedback was given to subjects indicating whether their decisions were correct or incorrect. This was done to maintain the subjects internal representation of the standard/base velocity. If the subject was responding to a stimulus travelling at the same speed as the standard then the feedback was random. Subjects responded by pressing the up/down arrows on a PC keyboard to signal whether the stimulus was going faster or slower than the standard.

In the training set, the velocity of each stimulus was predetermined. In the subsequent pair of interleaved runs, the velocity of each stimulus was determined by adaptive probit estimation (Watt & Andrews, 1981). Subjects were given the facility to repeat a stimulus, this facility was seldom utilised. The adaptive procedure was constrained such that the mean of the velocities of the test stimuli were equal to the velocity of the standard stimuli. With the first-order edge stimuli the polarity of the edge was randomly reversed across trials. All subjects were expert psychophysical observers with normal or corrected vision. Stimuli were viewed binocularly in a darkened room from a distance of two metres. Two slightly different forms of the same stimulus were presented simultaneously to the left and right of a fixation spot. Both of these images had the same stimulus speed, bar/edge width and type but were generated with different random values for spatial and temporal clipping. To try to minimise tracking one of the stimuli would be moving upwards whilst the other would be moving downwards. Subjects were instructed to maintain fixation on the fixation spot during stimulus presentation. Within experiments the order of presentation of the levels was randomised.

A psychometric function was fitted to the data from each run using probit analysis (Finney, 1971). The standard deviation of the subject's responses in terms of degrees per second represents the increment in velocity that would be needed for that subject to judge the test stimulus as moving faster than the standard stimulus 84% of the time. The usual criteria for velocity discrimination is 75%, a point that occurs 0.674 standard deviations from the mean on the normal distribution curve. The standard deviation was therefore multiplied by 0.674 to provide an estimate of the 75% point. This was divided by the velocity of the standard stimulus to yield the speed discrimination thresholds as a Weber fraction. Henceforth all thresholds are Weber fractions unless otherwise specified. For each stimulus level, for each subject, a minimum of 3 such measures were obtained. The graphs presented in the results section show the means and standard errors across these measures within each level. Data for each subject were subject to analysis of variance. Where appropriate, the results of these analyses are included in the text.

4.3. Results.

4.3.1. Speed discrimination thresholds as a function of velocity.

Within the two channel model, first- and second-order motion are detected through the operation of different sets of motion energy detectors. In the second-order channel these sets of detectors are preceded by filtering and non-linear stages which serve to demodulate the signal. Within each channel, an area of frequency space is sampled by a mosaic of motion energy detectors. Velocity may then be extracted by combining the outputs from a number of these detectors. Unless frequency space is sampled in exactly the same manner within both channels, we might reasonably expect that there may be some difference in the pattern of speed discrimination thresholds between first- and second-order stimuli over changes in the temporal parameters of stimuli. In the following pair of experiments, SDTs for first- and second-order bars and edges are measured as a function of velocity.

4.3.1.1. Experiment 4.1: Speed discrimination in first- and second-order bars as a function of velocity.

Three subjects were tested over 7 base velocities; 0.25, 0.5, 1.0, 2.0, 4.0, 8.0, and 16.0 degrees per second. For both the first- and second-order bars and over all stimulus levels, the standard deviation of the Gaussian was set to 0.9 arc minutes. Figure 4.4 shows Weber fractions for speed discrimination thresholds as a function of velocity. For all subjects SDTs for second-order bars are significantly higher than those for first-order bars (AJ, $F_{1,40} = 7.80$, $P < 0.01$; CB, $F_{1,45} = 9.53$, $P < 0.01$; WC, $F_{1,44} = 14.08$, $P < 0.01$).

For two of the subjects (CB & WC) the SDTs follow a U-shaped pattern with thresholds for first-order stimuli dropping to about 0.06 for a base velocity of $2.0°s^{-1}$. Whilst raised SDTs with low base velocities are a common finding (McKee, 1981; McKee *et al*, 1986; Panish, 1988); Orban *et al* (1985) found no upturn in thresholds until a velocity of about $64°s^{-1}$ was reached. In the present study the stimulus, when travelling at $16°s^{-1}$, traverses a maximum distance of 2.1° and a minimum distance of 1.0° (depending upon random spatial clipping). The stimulus will therefore only be moving for a period of between 60 and 130 ms. It seems reasonable to assume that the higher SDTs found at $16°s^{-1}$ are due to the brief amount of time that subjects have to extract velocity information from the stimulus (Snowden & Braddick, 1991). In the present study the stimuli were vertically moving bars of length 2.1° centred 1.4° right and left of the fixation spot. There is therefore the possibility that the fact that the stimulus was not centred on the fovea may be the root cause of the high SDTs associated with the velocity of $16°s^{-1}$. However, given that Orban *et al* (1985) found little effect on the upper end of the SDT curve with increasing eccentricity, the brief amount of time that this stimulus spent moving across the screen is still the best underlying cause of the comparatively high thresholds associated with this velocity.

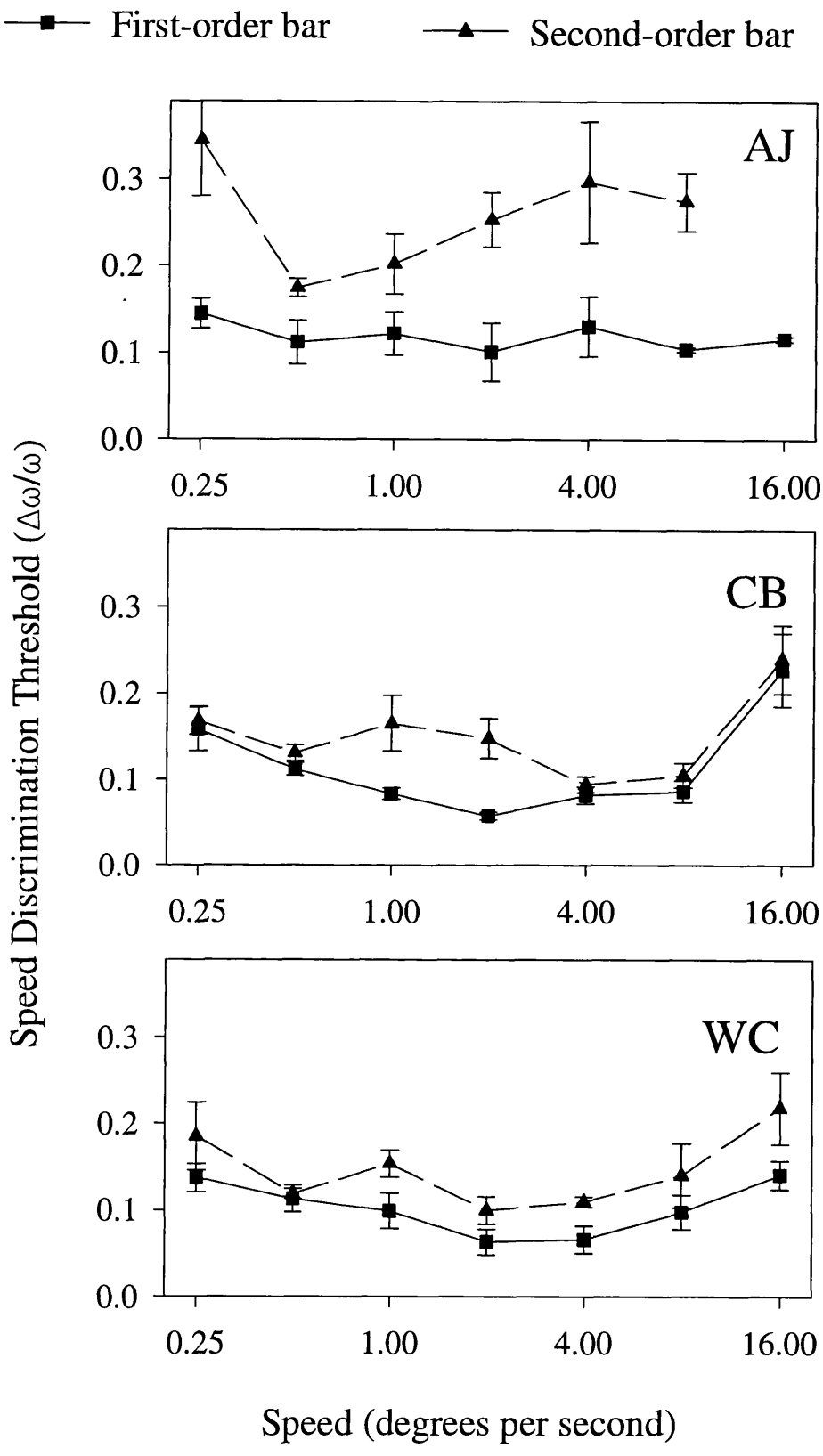


Figure 4.4: Weber fractions showing speed discrimination thresholds for first- and second-order Gaussian bars as a function of stimulus speed for three subjects. The bars indicate ±1 standard error.

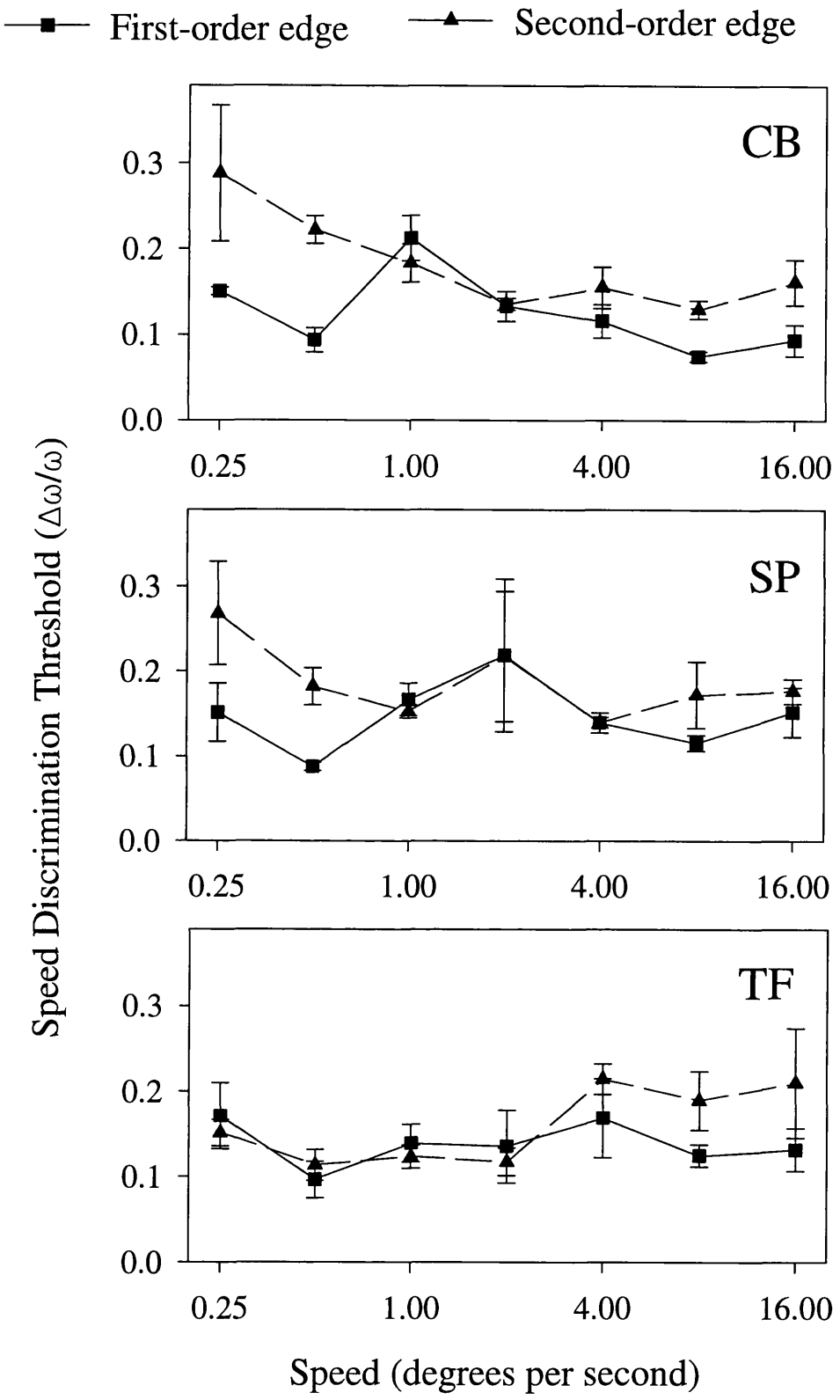


Figure 4.5: Weber fractions showing speed discrimination thresholds for first- and second-order Gaussian edges as a function of stimulus speed for three subjects. The bars indicate ±1 standard error.

4.3.1.2. Experiment 4.2: Speed discrimination in first- and second-order edges as a function of velocity.

Three subjects were tested over 7 base velocities; 0.25, 0.5, 1.0, 2.0, 4.0, 8.0, and 16.0 degrees per second. For both the first- and second-order edges over all stimulus levels, the width of the edge was 1.25 arc minutes. Figure 4.5 shows Weber fractions for speed discrimination thresholds as a function of velocity. Whilst thresholds are higher for second-order stimuli than for first-order stimuli at the higher and lower ends of the velocity range, this does not hold true for the mid-range velocities (1, 2 and 4 degrees per second). Over this range there appears to be little difference between the two types of stimulus. Results from the analysis of variance, showed that thresholds were significantly higher for the second-order stimuli for subject CB ($F_{1,35}$ = 16.73, $P < 0.01$), were marginally non-significant for subject SP ($F_{1,31}$ = 3.87, $P = 0.058$) and non-significant for subject TF ($F_{1,65}$ = 0.73, n.s.).

4.3.1.3. Discussion.

In terms of the shapes of the functions relating speed discrimination thresholds to stimulus velocity, there was no evidence for any systematic differences between first- and second-order stimuli. If first- and second-order velocity are extracted through separate motion energy mechanisms then not only must the method of operation of the two systems be very similar, but the way in which frequency space is sampled must also be similar. This observation is supported by data from Lu & Sperling (1995b) who proposed that the first- and second-order systems were characterised by the same temporal frequency response in terms of direction discrimination thresholds. Another possibility is that both first- and second-order motion are processed by a single mechanism.

In general, SDTs for second-order stimuli are higher than those for first-order stimuli. This difference was more pronounced with Gaussian bar stimuli than with the integral of Gaussian edge stimuli. If first- and second-order velocity are extracted through the operation of separate mechanisms, then this difference can be accounted for by proposing that the second-order system is simply less effective than the first-order system. However if first- and second-order velocity are extracted through the operation of a single system,

then the differences in the magnitudes of the SDTs must be accounted for in some other manner.

4.3.2. Speed discrimination thresholds as a function of width.

The previous two experiments looked at SDTs as a function of stimulus speed. It was proposed that a difference in the temporal frequency tuning of the first- and second-order systems would exhibit itself as a difference in the patterns of response to the two stimulus types over changes in the temporal parameters of the stimuli. No such evidence was obtained. However, given a two stage mechanism for velocity extraction, SDTs should also be dependent upon the spatial parameters of the stimuli. In the following pair of experiments, speed discrimination thresholds for first- and second-order bars and edges are measured across changes in the width of the stimuli.

4.3.2.1. Experiment 4.3: Speed discrimination in first- and second-order bars as a function of bar width.

Two subjects were tested on 5 bar widths with the standard deviations for the Gaussian set to the following: 0.45, 0.9, 1.8, 3.6 and 7.2 arc minutes. For both first- and second-order bars over all bar widths, the base velocity was set to $2.0°s^{-1}$. Figure 4.6 shows Weber fractions for speed discrimination thresholds as a function of bar width. The SDTs for the second-order bars are significantly higher than those for the first-order stimuli (CB, $F_{1,22}$, $P < 0.01$; WC, $F_{1,22}$, $P < 0.01$). For both types of stimulus there is no significant effect of bar width on SDTs with thresholds around 0.05/0.06 for the first-order bars and around 0.09 for the second-order bars. Compared to experiment 4.1 the SDTs found in the present experiment are slightly lower. A possible explanation is that whilst base velocity was varied in experiment 4.1, it was always kept at $2°s^{-1}$ in this experiment. It may well be easier for subjects to hone their ability to discriminate around a single base velocity without the distraction of having to discriminate other base velocities.

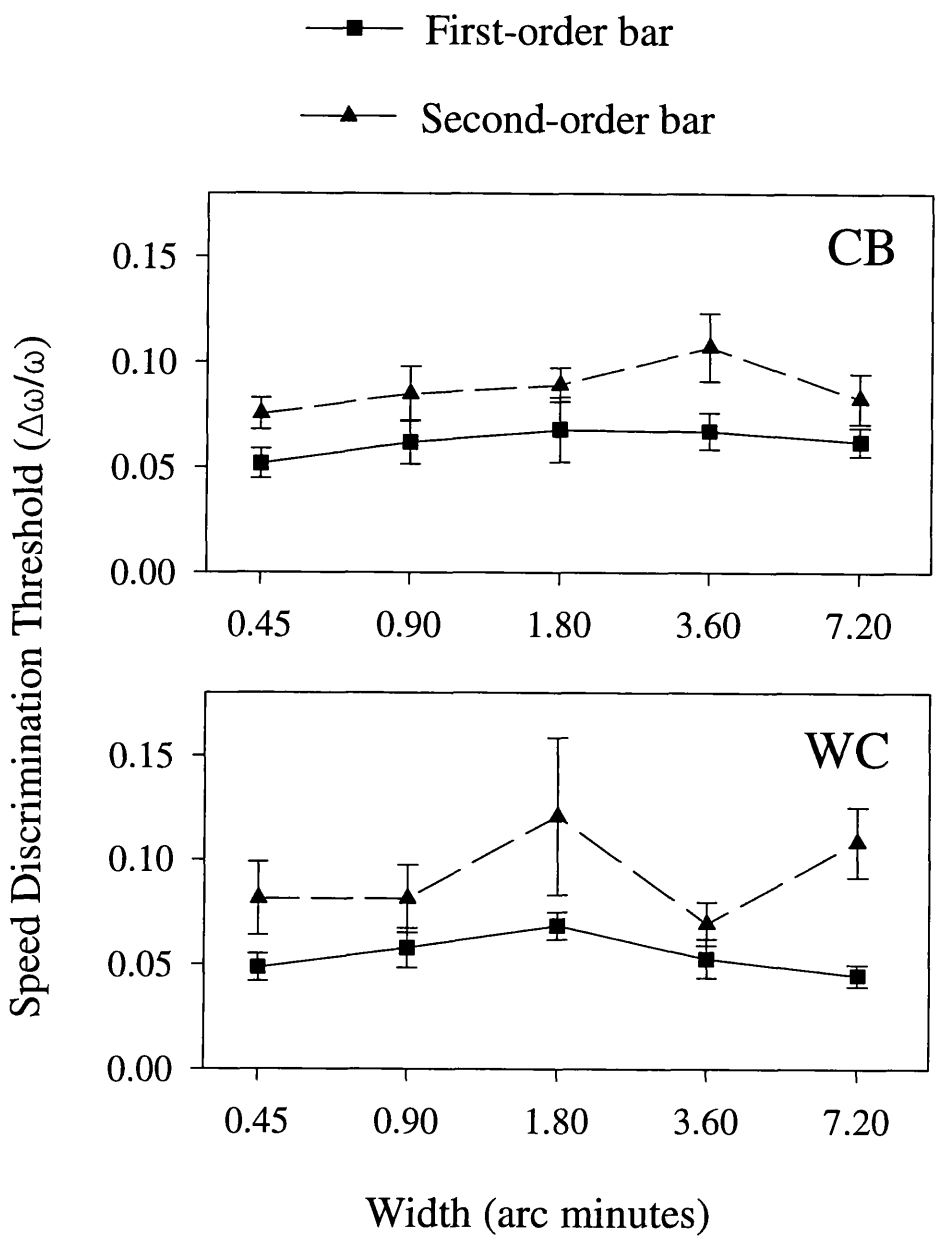


Figure 4.6: Weber fractions showing speed discrimination thresholds for first- and second-order Gaussian bars as a function of stimulus width for two subjects. The bars indicate ±1 standard error.

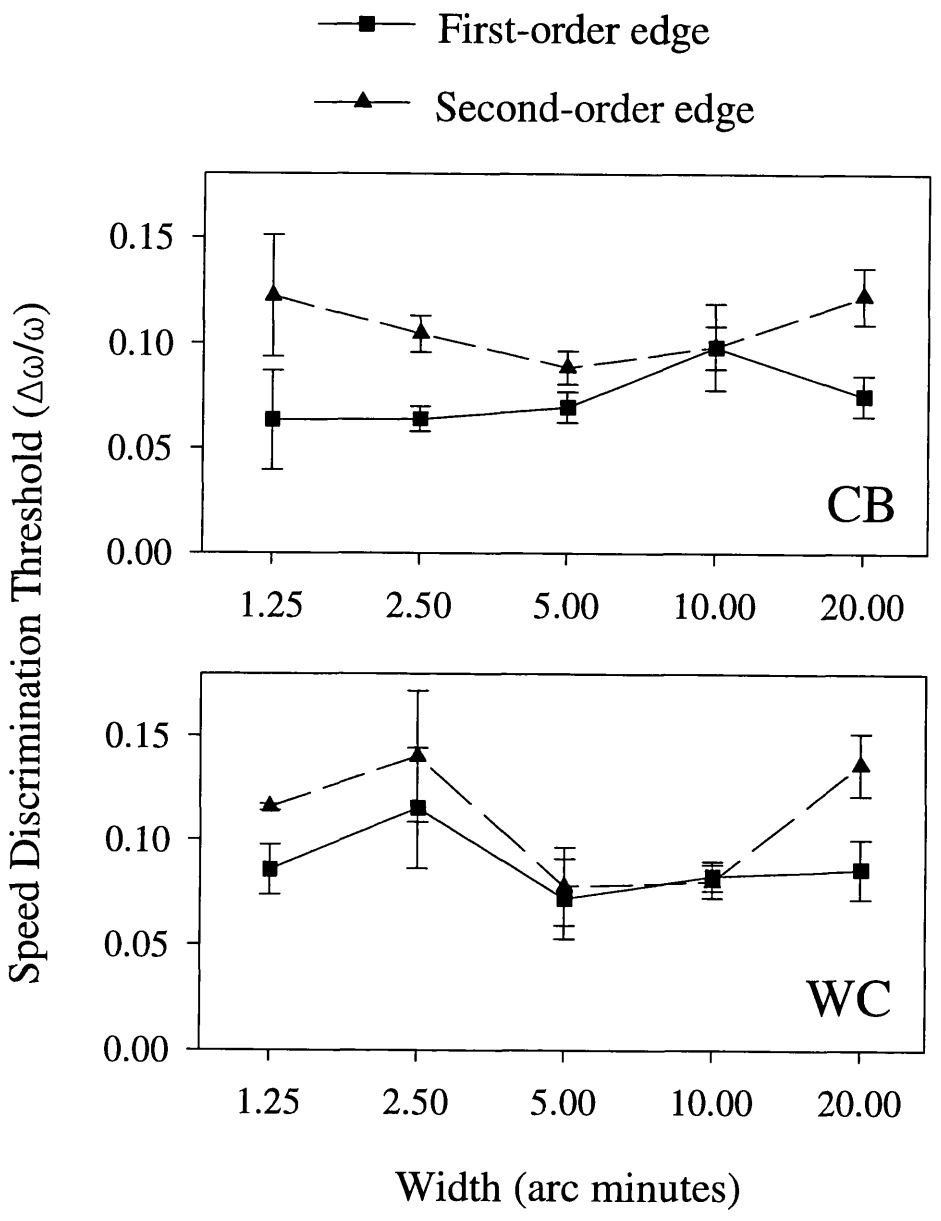


Figure 4.7: Weber fractions showing speed discrimination thresholds for first- and second-order Gaussian edges as a function of stimulus width for two subjects. The bars indicate ±1 standard error.

4.3.2.2. Experiment 4.4: Speed discrimination in first- and second-order edges as a function of edge width.

Two subjects were tested over the following 5 edge widths: 1.25, 2.5, 5, 10 and 20 arc minutes. The standard velocity was set $2.0°s^{-1}$. Figure 4.7 shows Weber fractions for speed discrimination thresholds as a function of edge width. Generally, SDTs appear to be slightly higher for the second-order stimuli but this is not true across all widths tested. Analysis of variance revealed no significant difference between the stimulus types (CB, $F_{1,23} = 3.61$, $P = 0.07$; WC, $F_{1,22} = 3.7$, $P = 0.067$). With edges of width 10 arc minutes, both subjects perform equally well on both the first- and second-order edges. Compared to the data from experiment 4.2, it can be seen that the SDTs found in this experiment are lower and are more comparable with the figures for first-order stimuli from other studies. Furthermore, for both subjects in the present experiment, SDTs were lower for first-order edges than for second-order edges at a stimulus width of 1.25 arc minutes (the width of the edges in experiment 4.2).

4.3.2.3. Discussion.

The results provide no evidence for a difference in spatial frequency tuning between first- and second-order mechanisms. Over all 4 stimuli studied, changes in the width of the stimuli appear to make little difference to the thresholds. The observations concerning the relative magnitudes of the first- and second-order stimuli from the previous pair of experiments are supported. SDTs in response to second-order bars were significantly higher than those for first-order bars. With first- and second-order edges the results of both subjects approached, but did not reach, significance. Although the pattern of difference in magnitude appears clear across the four experiments described above, it might be argued that the contrast of the stimuli needs to be normalized with respect to detection thresholds for a true picture of the relationship to occur.

4.3.3. Speed discrimination thresholds as a function of contrast.

McKee *et al* (1986) showed that velocity discrimination in sinusoidal gratings was independent of contrast (when contrast was varied from 0.05 to 0.82) and Turano & Pantle (1989) found that contrast needed to be reduced to about 0.05 for a significant reduction in performance in speed discrimination on an apparent motion sine wave grating. Turano & Pantle (1989) also investigated velocity discrimination in a second-order stimulus (contrast-modulated gratings) and showed that velocity discrimination thresholds remained constant between modulation depths of 0.6 and 1.0. However, Cropper (1994) found that SDTs for second-order stimuli (beats) approached optimal performance more slowly that those for luminance gratings when expressed in terms of multiples of their respective contrast detection thresholds. In order to test whether the stimuli examined in this chapter have reached saturation in terms of speed discrimination thresholds, the following two experiments examine SDTs in first- and second-order bars and edges as a function of the amplitude of the stimuli.

4.3.3.1. Experiment 4.5: Speed discrimination in first- and second-order bars as a function of amplitude.

The first- and second-order bars may be regarded as modulations of luminance and contrast respectively. In this experiment amplitude is measured as a proportion of the amplitudes described in the methods section. For example the first-order bar varies from mean luminance (15.5 cd/m^2) to 30 cd/m^2, giving a difference of 14.5 cd/m^2 between the two extremes. When the amplitude of the first-order bar is 0.4 then this difference is reduced to four tenths of that value (ie. 5.8 cd/m^2) so that the peak luminance of the bar would be 21.3 cd/m^2. In the case of the second-order bar, the function varies between a contrast of zero and a contrast of 0.94. When the amplitude is 0.4 then this maximum contrast is reduced to four tenths of its value (ie. 0.376).

Two subjects were tested over three modulant amplitudes: 0.2, 0.4 and 0.8. The data are shown in Figure 4.8. It is clear that there is little or no effect of amplitude. In both subjects, the thresholds for the second-order stimuli were significantly higher than those for the first-order stimuli (CB, $F_{1,14} = 22.32$, $P < 0.01$; AJ, $F_{1,12} = 17.69$, $P < 0.01$).

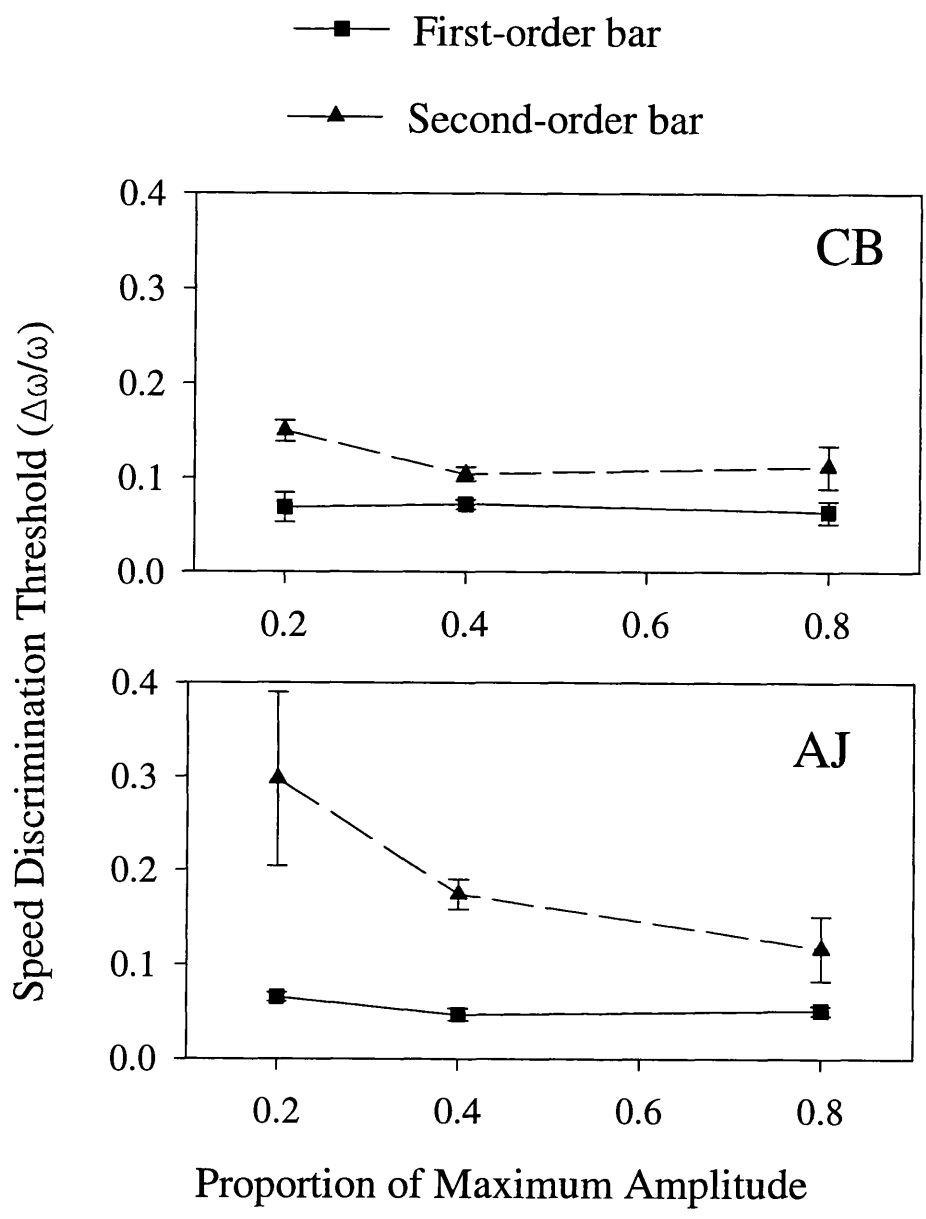


Figure 4.8: Weber fractions showing speed discrimination thresholds for first- and second-order Gaussian bars as a function of a proportion of maximum stimulus amplitude. The bars indicate ±1 standard error.

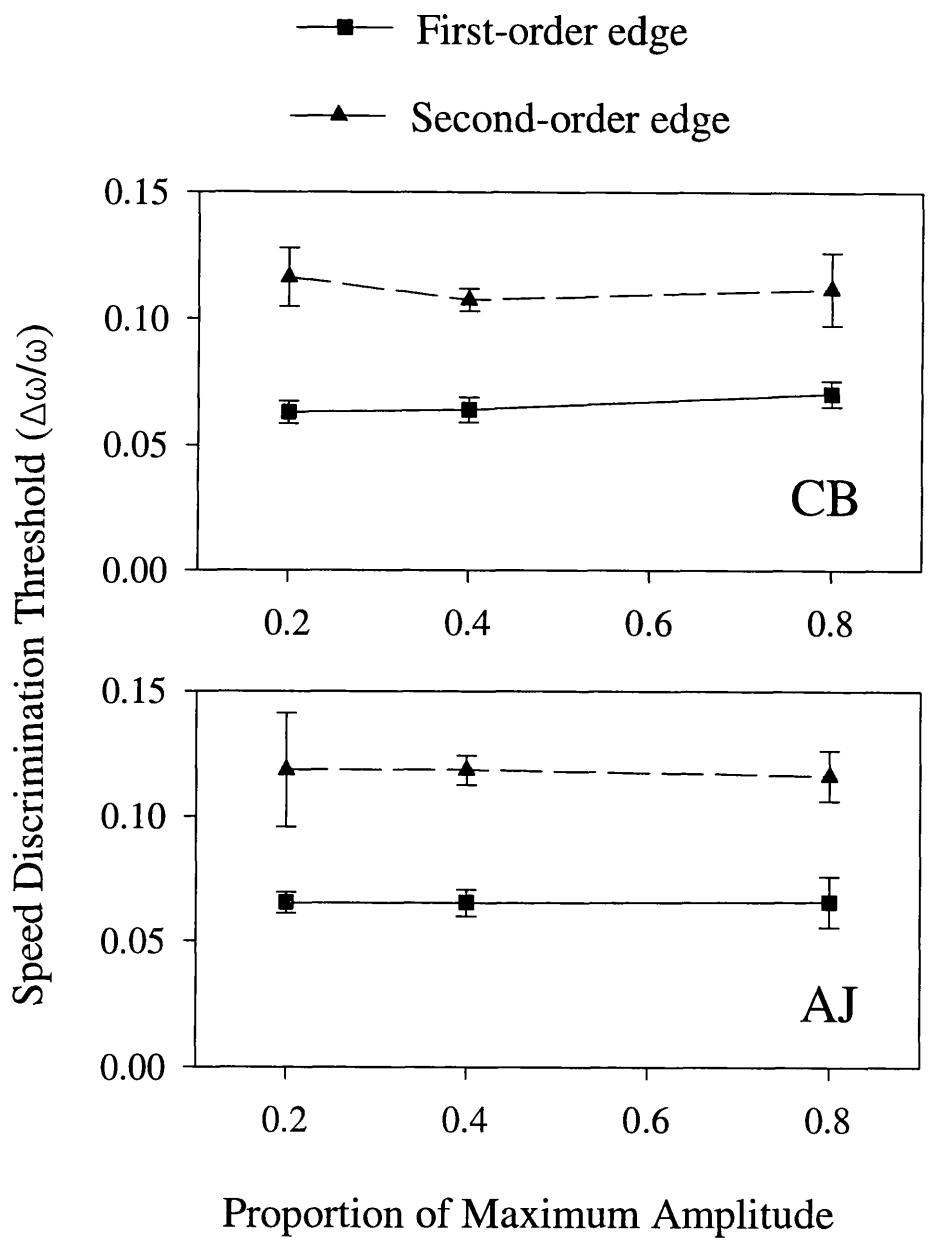


Figure 4.9: Weber fractions showing speed discrimination thresholds for first- and second-order Gaussian edges as a function of a proportion of maximum stimulus amplitude. The bars indicate ±1 standard error.

4.3.3.2. Experiment 4.6: Speed discrimination in first and second-order edges as a function of amplitude.

As in experiment 4.5, amplitude is measured as a proportion of the amplitudes described in the methods section. For example the first-order edge varies from a minimum of 1.0 cd/m^2 to a maximum of 30 cd/m^2, a difference of 29 cd/m^2. When the amplitude is 0.4 then this difference is reduced by a factor of four tenths to give 11.6 cd/m^2. The minimum luminance becomes 9.7 cd/m^2 and the maximum luminance becomes 21.3 cd/m^2 (the mean of the maximum and minimum luminances is always mean luminance). In the case of the second-order edge, the functions runs between a maximum contrast of +0.94 and a minimum contrast of -0.94. The contrast is simply reduced by the relevant factor.

Two subject were tested over three modulant amplitudes, 0.2, 0.4 and 0.8. The data are shown in Figure 4.9. There is no effect of amplitude. Thresholds for second-order stimuli are significantly higher than those for first-order stimuli (CB, $F_{1,17} = 33.13$, $P < 0.01$; AJ, $F_{1,15} = 13.13$, $P < 0.01$).

4.3.3.3. Discussion.

In the context of a speed discrimination task, scaling with respect to detection thresholds involves at least two assumptions. Firstly, that the stimulus is detected by the same mechanism(s) involved in the measurement of speed. Secondly, that the efficiency of whatever mechanisms underlie speed discrimination scale similarly with stimulus contrast for different stimuli. Certainly, within the motion from Fourier components framework, it is not accepted that detection involves exactly the same mechanisms as those necessary to gain some measure of speed. Even the simplest method of velocity extraction involves the operation of at least two detectors to remove the confounding effect of stimulus contrast (Adelson & Bergen, 1985). Indeed, there is no scheme in the current literature where the measurement of velocity is gained from the output of a single motion energy detector. However the detection of a stimulus could well be based upon the output of a single spatio-temporal filter. It would therefore seem that the first assumption is necessarily false. The second assumption is itself based upon assumptions as to the linearity of the system with

respect to the components of different stimuli and, more critically, with the way in which the components are combined.

It was precisely because of the objections raised above that, in the present chapter, no attempt was made to scale modulant amplitudes with respect to detection thresholds. Rather, the amplitude of the modulants was set to the maximum value that could be obtained on the system so that the experiments measure the response of the system at saturation. The results from experiments 4.5 and 4.6 show that the stimuli employed in experiments 4.1 to 4.4 are well above levels where the response of mechanisms involved in the perception of velocity have reached saturation with respect to stimulus amplitude.

4.4. General Discussion.

Speed discrimination thresholds for first- and second-order bars and edges were compared over velocity, width and stimulus magnitude. Over the spatial and temporal manipulations there was no evidence for any difference in the patterns of response for first- and second-order stimuli. However, there were statistically significant differences between the magnitudes of the SDTs obtained in response to these stimuli, with SDTs for the second-order stimuli higher than those for the first-order stimuli. This was more pronounced with first- and second-order bars than with first- and second-order edges.

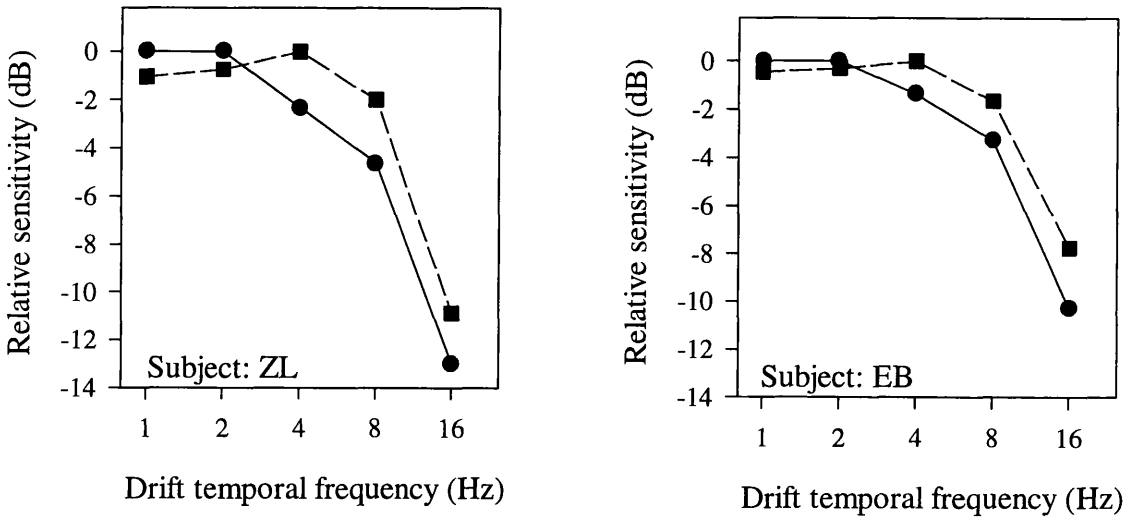
It is possible that measures of contrast sensitivity may relate to supra-threshold measures of performance. For example, there has been some debate on the relationship between perceived velocity and contrast detection thresholds for pattern and flicker (Harris, 1980; Green, 1981; Burbeck, 1981), and it has been proposed that velocity is encoded as the ratio of the ratio of the outputs of flicker and pattern channels (Smith, 1985). This ties in with Adelson & Bergen's (1985) model whereby the output of a directionally selective opponent energy stage is contrast normalized by dividing the output by that of a static energy detector. If the extraction of velocity is based upon an initial directional discrimination stage, then it seems reasonable to assume that, as contrast thresholds for direction discrimination increase or decrease over spatial and/or temporal frequency, so too should speed discrimination thresholds. However, it should be noted that this is an assumption and is not necessarily true. Measures of contrast thresholds are measures of the sensitivity of the system whereas

speed discrimination thresholds, at least in the two stage model described above, are dependent upon the density of the sampling of frequency space.

The pattern of data concerning direction discrimination over temporal frequency in second-order stimuli is far from clear. Lu & Sperling (1995b) examined direction detection thresholds for translating sine wave gratings and contrast modulations of static noise and found no difference in temporal frequency response between the two stimulus types. However, Smith & Ledgeway (1997b) compared direction discrimination in contrast modulations of static and dynamic noise. They found that the pattern of response over temporal frequency appeared bandpass in the case of the contrast modulation of static noise, but the pattern of response appeared lowpass in the case of the dynamic noise carrier. The data from Lu & Sperling (1995b) and Smith & Ledgeway (1997b) are re-plotted in Figure 4.10. Note that the scales of the axes are identical between the graphs and that for the contrast modulations of static noise the results of the two studies are remarkably similar. In a study that examined direction detection thresholds in beats and sine wave gratings, Holliday & Anderson (1994) showed that the peak response for beat patterns occurred at a lower temporal frequency than that obtained for sine wave gratings (~1-2 Hz compared to ~10 Hz). Additionally, the pattern of response to beats appeared bimodal with a second smaller peak occurring at ~10 Hz. There therefore appears to be no clear relationship between temporal frequency and direction discrimination thresholds in second-order stimuli in general. To a large degree the results appear to be based on the particular second-order stimulus employed.

There is little data comparing direction discrimination thresholds in first- and second-order stimuli over modulant spatial frequency. In general, studies that examine motion stimuli investigate the response over temporal frequency whilst those that examine static stimuli investigate the response over spatial frequency. However, there is some agreement that the second-order system is biased towards lower modulant spatial frequencies than the first-order system. This is incorporated into Wilson *et al*'s (1992) two channel model. In one study that did examine motion detection thresholds for first- and second-order stimuli as a function of modulant spatial frequency, the response curve appeared lowpass for sinusoidal contrast modulations of static noise and bandpass for sine wave gratings (Nishida, Ledgeway & Edwards, 1997). The notion of lower spatial frequency selectivity for second-

a) Lu & Sperling (1995b)



b) Smith & Ledgeway (1997b)

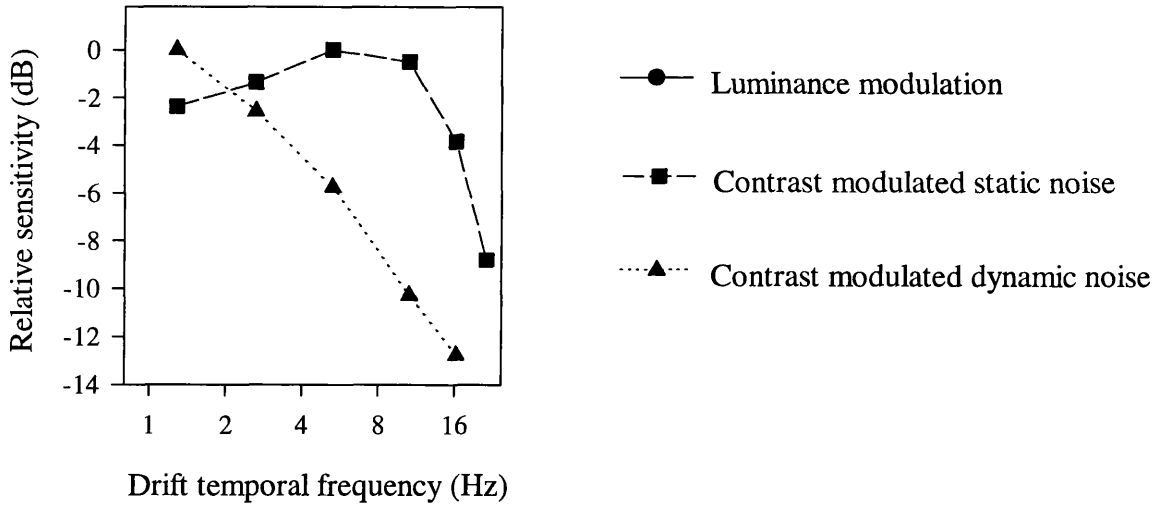


Figure 4.10: Contrast thresholds for direction detection plotted in terms of relative sensitivity as a function of temporal frequency. Data re-plotted from (a) Lu & Sperling, 1995b and (b) Smith & Ledgeway, 1997b.

order motion is also supported by Jamar & Koenderink (1985) who showed that modulation detection thresholds for contrast modulated noise gratings showed a lowpass response over spatial frequency. Detection thresholds for a sine wave grating showed a bandpass response with the lowest contrast thresholds at ~3 cycles per degree.

In terms of the patterns of results described in this chapter, there is no evidence for any difference over spatial and temporal parameters between the first- and second-order motion stimuli. In order to support the two channel model one would have to suppose that the two channels are alike enough so that there appear to be no differences in terms of the pattern of their SDTs, but that the channels may possibly differ from one another at some stage that precedes velocity extraction. This latter point is necessary to account for any differences in response to other measures of the sensitivity of the systems that may occur, particularly any degree of independence between speed discrimination thresholds and contrast thresholds for direction discrimination that may occur. For example Cropper (1994), examining the perception of motion in beat stimuli, found that, with increases in beat contrast, speed discrimination thresholds dropped to a level compatible with those found for luminance gratings. He notes that this pattern is not seen with direction discrimination.

Of course the lack of a difference in the patterns of response also supports the hypothesis that first- and second-order velocity are extracted by the same mechanism. If first- and second-order motion are detected by a single mechanism then the differences in magnitudes of the speed discrimination thresholds between the first- and second-order stimuli must adequately be accounted for. SDTs for second-order stimuli were higher than those for first-order stimuli. This difference was more pronounced with Gaussian bar stimuli than with the integral of Gaussian edge stimuli. Given the results of work presented in Chapter 3 it may be profitable to suppose that the larger first-order/second-order differences for bar stimuli compared to edge stimuli may be due to the different carrier types. The second-order bar is a contrast modulation of dynamic noise, the second-order edge stimulus is a modulation of static noise.

Figure 4.11 shows Fourier spectra for the contrast-defined stimuli. These were created by adding the Fourier amplitude spectra of 100 instantiations of each of the respective images. It is clear that, in the case of the second-order edge, far more of the energy in the image lies

on the line signifying components with a temporal frequency of zero than in the case of the second-order bar. Therefore in the latter, a greater proportion of the expected first-order energy in the image will contain motion information than in the case of the second-order edge. It would seem reasonable to suppose that an increased proportion of random first-order motion energy would lead to increased thresholds for speed discrimination. This assertion is supported by evidence from Bowne, McKee & Glaser (1989) who showed that, in a two-dot apparent motion task, speed discrimination thresholds were raised by the inclusion of elements that added additional motion signals into the local velocity field.

The data presented in this chapter therefore supports two hypotheses. Firstly, that first- and second-order velocity are measured by different but very similar systems. It is not clear how this hypothesis can account for any differences that do exist in the perception of first- and second-order stimuli. Secondly; that first- and second-order velocity are extracted through the same mechanism. In the latter, any differences that do arise between the perception of motion elicited by the two stimulus types must be explained by appealing to differences in the form of the stimuli rather than upon the operations of separate mechanisms.

4.5. Conclusion.

There were no systematic differences in the patterns of results between the first- and second-order stimuli over velocity and width. This implies that either separate but very similar mechanisms measure velocity in both types of stimulus or that first- and second-order velocity are extracted by the same mechanism. The latter of these two options presents the simplest explanation of the results. There were systematic differences in the magnitudes of the speed discrimination thresholds. Those for the second-order stimuli were higher than those for the equivalent first-order stimuli. This was more pronounced with first- and second-order bars than with first- and second-order edges. It is possible that these differences may be accounted for by appealing to the presence of the different carriers in the second-order stimuli.

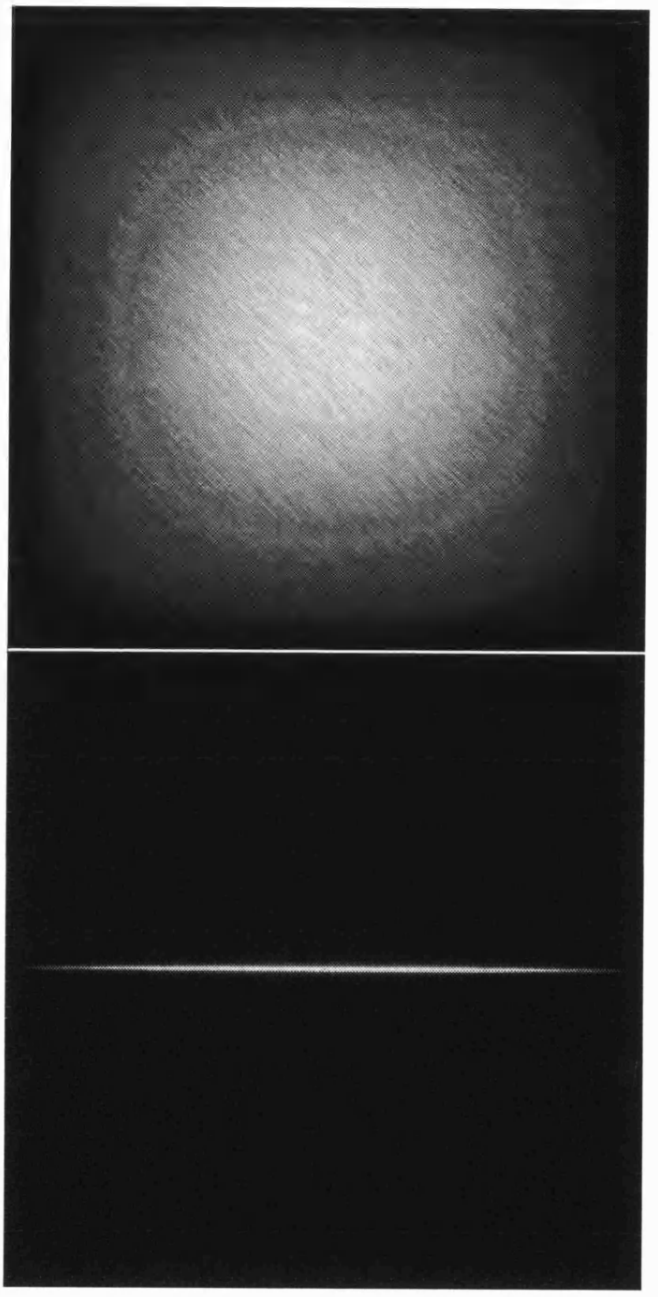


Figure 4.11: Fourier spectra created from the addition of 100 instantiations of a second-order Gaussian bar (top) and 100 instantiations of a second-order Gaussian edge (bottom).

Chapter Five

Perception of motion direction in luminance and contrast defined reversed-phi motion sequences.

5. Perception of motion direction in luminance and contrast defined reversed-phi motion sequences.

Non-linear processing can be used to recover the motion of contrast modulations of binary noise patterns. A non-linear stage has also been proposed to explain the perception of forward motion in motion sequences which typically elicit reversed phi. In this chapter, the perception of motion direction in luminance and contrast defined reversed-phi stimuli was investigated. A percept of forward motion could be elicited by both types of stimuli. The perceived direction of motion seen in the contrast defined stimuli showed a profound carrier dependency. The replacement of a static carrier by a dynamic carrier can reverse the perceived direction of motion. Forward motion was never seen with dynamic noise carriers. For luminance and contrast defined patterns the reversed motion percept increasingly dominated with increases in the spatial frequency and temporal frequency of the modulation. Differences in the patterns of responses to the two stimuli over spatial and temporal frequency were abolished by the *addition* of noise to the luminance defined stimulus. It is argued that these data suggest the possibility that a single mechanism may mediate the perception of both luminance and contrast defined motion.

5.1. Introduction.

Reversed apparent motion was first identified and investigated by Anstis (1970). If, in a two frame apparent motion sequence, the contrast of the second frame is reversed, then a strong percept of motion is obtained in the direction opposite to the displacement. This finding has generally been taken as evidence supporting the hypothesis that human motion extraction is based on motion energy analysis. Chubb & Sperling (1989a) describe a reversed-phi stimulus (which they refer to as Γ) which consists of a cyclic pattern of moving contrast reversing bars. These bars step forward a quarter of a cycle every frame and reverse their contrast (see Figure 5.1a). Chubb & Sperling report that subjects viewing the stimulus see motion in the *forward* (direct-phi) direction when close to the stimulus but see motion in the *reversed* (reversed-phi) direction when they move further from the stimulus. This reversal (from forward to reversed motion) has also been shown to be elicited by increases in spatial frequency and temporal frequency of the stimulus (Gorea, 1995).

These findings are of particular interest because, within the framework of motion from Fourier components, a percept of forward motion can only reasonably be explained by supposing that some rectifying non-linearity is applied around mean luminance prior to motion energy analysis. Chubb & Sperling therefore propose that this phenomenon is indicative of a two channel architecture in low level vision. The reversed motion percept is identified with the linear channel and the forward motion percept with a rectifying second-order channel. It should however be noted that this change in perceived direction can also be interpreted as resulting from the scale-dependent behaviour of a single motion mechanism (Johnston *et al*, 1992; Johnston & Clifford, 1995b).

Chubb & Sperling (1988) describe a class of stimuli which they refer to as micro-balanced; contrast modulated random noise provides an example of this type of stimulus. They show that standard motion detectors will show no systematic response to a micro-balanced stimulus, and therefore any perceived motion in such a stimulus must be due to activity in some non-Fourier channel. If a reversed-phi stimulus were used to modulate the contrast of a random noise field (see Figure 5.1c), then a two channel model would predict that no motion energy would be detected by the Fourier channel. The processing in the second-order channel should result in the recovery of the modulation. Therefore motion detectors in the second order channel should signal reversed motion which we can refer to as second-order reversed-phi. If forward motion were elicited by this stimulus then this would need to be explained either by postulating an additional rectification stage or by supposing that some additional mechanism, based upon a different set of principles, might exist.

Reversed apparent motion has been shown to occur with contrast reversing Random Dot Kinematograms (Sato, 1989). Using the second-order equivalent of this stimulus, a "Random Window Kinematogram", Nishida (1993) demonstrated that reversed-phi can be elicited by a non-Fourier stimulus. Nishida took this finding as evidence for motion energy analysis within the second-order system. With the second-order analogue of a "four stroke" reversed motion stimulus studied by Anstis & Rogers (1986), Mather & Murdoch (1996) showed that continuous reversed motion could be evoked by a second-order stimulus. Two further studies (Sperling & Lu, 1996 and Benton, Johnston & McOwan, 1996) used contrast defined versions of the Chubb & Sperling (1989a) stimulus. Whilst both studies found second-order reversed phi with this stimulus, Sperling and Lu additionally found that motion

in the forward direction could also be elicited when stimuli were viewed in foveal vision at low temporal frequencies. In a similar vein to the approach taken by Chubb & Sperling (1989a), Sperling & Lu explain the motion reversal as the result of two motion systems. The second-order channel is identified as that which detects reversed motion whilst the feature tracking mechanism is seen as that which detects forward motion.

It is clear then that the investigation of the motion elicited by first-order, and more recently second-order, reversed-phi sequences has played an important part in attempts to elucidate the mechanisms by which we perceive motion. In this chapter the perception of motion direction in these sequences is systematically investigated over a range of spatial and temporal frequencies. The effects of manipulating stimulus magnitude and eccentricity are also examined. Second-order motion mechanisms seek to extract the motion of the modulant whilst discarding the carrier. If the mechanism is efficient then such an approach should predict little effect of carrier type on the perception of second-order motion. However recent studies have show that the form of the carrier can affect the perception of second-order motion (ie. Smith & Ledgeway, 1997a). Therefore, in this study, the effects of any interactions between carrier type and the various stimulus dimensions are explored. Furthermore, the effects of additive noise on the perception of motion direction in first-order reversed-phi sequences are also investigated.

5.2. Method.

5.2.1. Equipment.

Images were displayed on a non-interlaced monochrome Manitron monitor, driven by a Matrox Image-1280 graphics card controlled by an IBM compatible PC. The graphics card delivered 8 bits per pixel to give 256 grey levels. The display was carefully gamma corrected using a Minolta LS-110 luminance meter and was spot checked using a UDT OPTOMETER S370. All of the sequences that were generated and displayed were 32 frames long. The frame rate was 59.5 Hz. In each trial the screen around the stimulus was set to mean luminance and the full screen reverted to mean luminance between trials. Stimuli were displayed in a vertically oriented rectangle positioned in the centre of the screen. From a viewing distance of 1m the screen had a width of 20.80° and a height of 16.67°. The stimuli

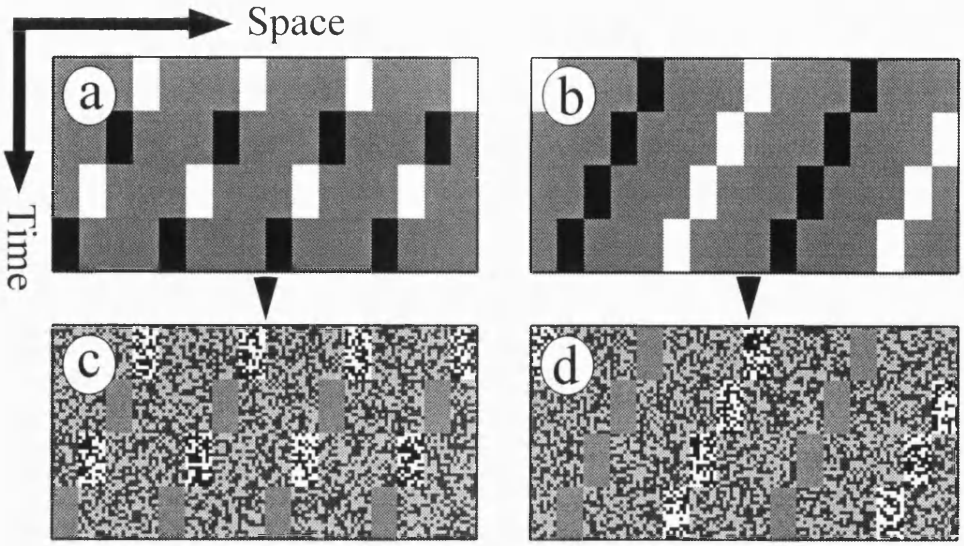


Figure 5.1: Space-time plots showing: (a) Γ_L stimulus; (b) Φ_L stimulus; (c) Γ_C stimulus; (d) Φ_C stimulus.

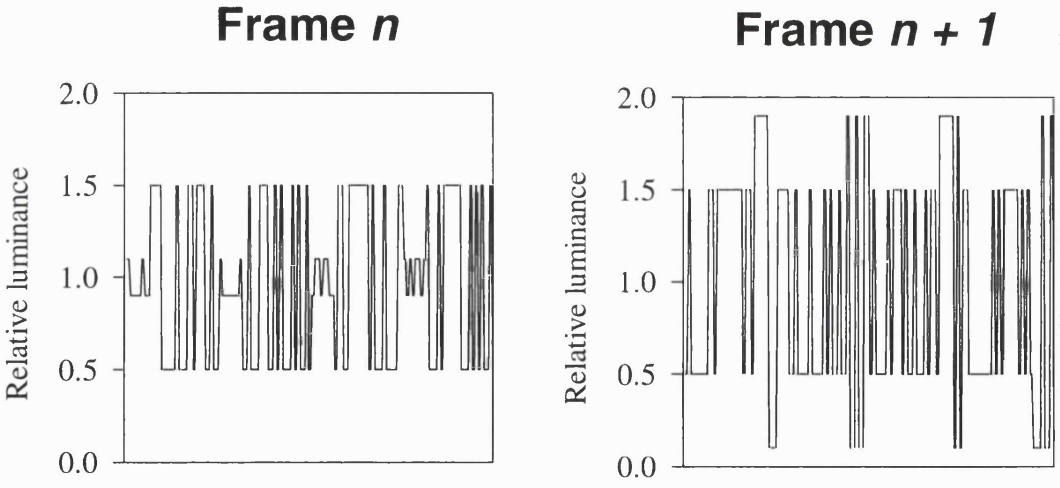


Figure 5.2: Luminance cross-sections from successive frames of a Γ_c stimulus with a dynamic binary noise carrier. Luminance is given in multiples of mean luminance.

stepped forward ¼ cycle every stimulus frame. The temporal frequency of the stimulus is dependent upon the number of physical frames per stimulus frame. For example, if there are 4 physical frames per stimulus frame then the temporal frequency is 3.75 Hz. The direction of displacement was either upwards or downwards.

Experiment 5.1: Gamma corrected space-time images were constructed in PC RAM and passed to the graphics card. Data was loaded, frame by frame, from the space-time image into an output look-up table, which was indexed by a ramp drawn in display memory. The image was displayed in a rectangle centred in the middle of the screen. From a viewing distance of 1m the rectangle had a width of 12.55° and a height of 14.59°. Mean luminance (I_0) was 15.4 cd/m^2.

Experiments 5.2 to 5.8: Gamma corrected frames were generated in PC RAM and stored on the graphics card. The rectangle within which stimuli were displayed had a width of 8.40° and a height of 12.54°. Mean luminance (I_0) was 14.8 cd/m^2. As many of the image sequences in these experiments consisted of a series of two dimensional spatial patterns, far more information has to be stored on the graphics card than is the case when one dimensional images are used (ie. experiment 5.1). Because of the limited amount of graphics memory available, the "images" on the graphics card were scaled up 4 times for display on the screen. Each "stimulus pixel" therefore consists of 4x4 "physical pixels", where a physical pixel is one of the 1280x1024 picture elements available on the monitor.

5.2.2. Stimuli.

Luminance defined stimuli: Two basic sets of stimuli were constructed: a polarity reversing bar stimulus, described as "Γ" by Chubb & Sperling (1989a) and shown in Figure 5.1a, and a comparison stimulus. This comparison stimulus, a Φ motion sequence, is shown in Figure 5.1b and is a set of light and dark forward stepping bars. Both of these stimuli step forward ¼ cycle each stimulus frame. The images consist of three intensity levels, I_{max}, I_0, and I_{min}, where I_0 is the mean luminance. The relationships between the three luminance levels may be described in the following way:

$$I_0 = (I_{max} + I_{min}) / 2 \qquad \text{(where } I_{max} \geq I_{min}) \ .$$

Stimulus contrast is defined by the following equation:

$$\text{Contrast} = (I_{max} - I_{min}) / (I_{max} + I_{min}) .$$

Contrast defined stimuli: The envelopes used to modulate luminance were also used to modulate the contrast of various types of binary random noise. This was done such that, in the subsequent image, areas of maximum luminance became areas of maximum contrast (C_{max}), areas of medium luminance became areas of medium contrast (C_0), and areas of low luminance became areas of low contrast (C_{min}). Each contrast region consisted of two luminance levels as detailed below

$$C_{max}: \; I_0 \pm k_{max} , \qquad C_0: \; I_0 \pm k_0 , \qquad C_{min}: \; I_0 \pm k_{min} ,$$

where

$$k_0 = (k_{max} + k_{min}) / 2 , \qquad (k_{max} \geq k_{min}) .$$

The standard Michelson contrast of each region is given by

$$C_x = k_x / I_0 .$$

The relationship between the three contrast levels may be described by

$$C_0 = (C_{max} + C_{min}) / 2 \qquad (\text{where } C_{max} \geq C_{min} \text{ and } C_0 = 0.5) .$$

For the purposes of this chapter, modulation depth is defined as

$$\text{Modulation Depth} = (C_{max} - C_{min}) / (C_{max} + C_{min}) .$$

Space-time plots of the contrast defined stimuli are shown in Figures 5.1c and 5.1d; cross sections of the luminance profiles from two frames of the contrast defined Γ stimulus are

shown in Figure 5.2. Given the equations above and the mean luminance (I_0) used in the experiments, the majority of the stimuli used this chapter can be characterised by their contrast (in the case of luminance defined stimuli) or modulation depth (in the case of contrast defined stimuli). The exceptions occur when stimuli contain additive noise. In these cases, details of the stimulus before noise has been added and the amplitude of the noise provide a characterisation of the stimulus. Where it is not obvious from the context, contrast defined stimuli are marked as such by a subscripted "C" (ie Φ_C and Γ_C). Luminance defined stimuli are indicated by a subscripted "L" (ie Φ_L and Γ_L).

Binary random noise: The contrast defined stimuli used in this chapter are modulations of binary random noise. Four types of binary random noise are used: static one dimensional (1D) noise, dynamic 1D noise, static 2D noise and dynamic 2D noise. Figure 5.3 shows four frames from Γ_C stimuli with these four carrier types.

5.2.3. Procedure.

All subjects had normal or corrected to normal vision. Subjects were asked to fixate in the centre of the monitor screen and judge whether stimuli were moving upwards or downwards. The middle of the screen was approximately at eye level and the screen was the only source of illumination. There was a minimum gap of 1 second between trials. In all experiments, subjects were presented with a number of different stimulus types. In all but experiment 5.1, these stimulus types were randomly interleaved. In experiment 5.1 the stimuli were presented in blocks of one type; the order of presentation of the blocks was randomized across subjects. The direction of displacement and the start position of the bars were selected at random. In the Γ_L and Γ_C stimuli, the initial polarity of the modulant bar was also randomized. Effectively, all stimuli were stacked within an experiment, and a stimulus was randomly chosen from this stack for each trial. After a judgement had been made the stimulus was removed from the stack. Subjects indicated the perceived direction of motion using the up/down arrows on a PC keyboard. If subjects suffered a lapse of attention on a particular trial they could restack the stimulus which they had just been shown by means of a key press. A fresh stimulus would then be randomly selected. This facility was seldom used.

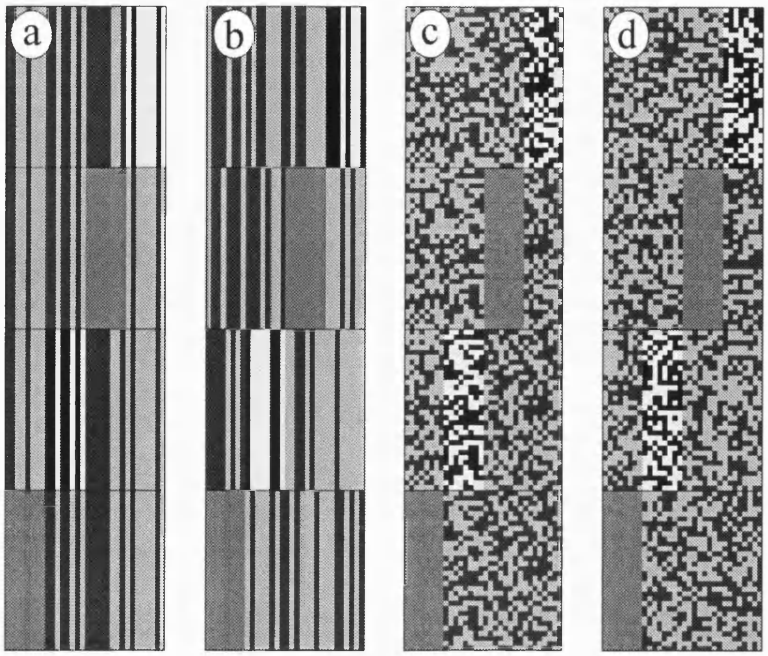


Figure 5.3: Four successive frames from a Γ_L stimulus with the following noise carriers: (a) 1D static noise, (b) 1D dynamic noise, (c) 2D static noise, (d) 2D dynamic noise.

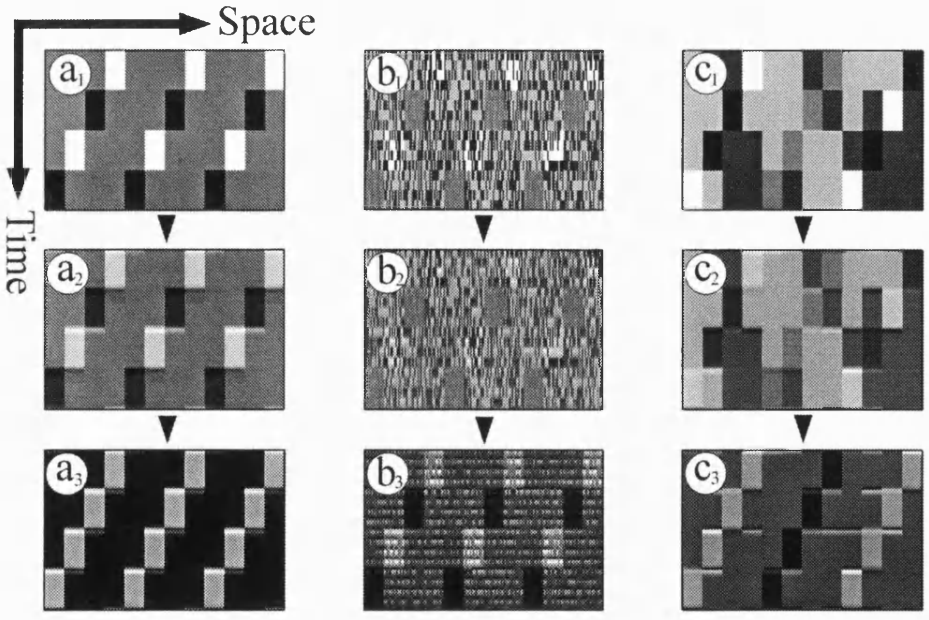
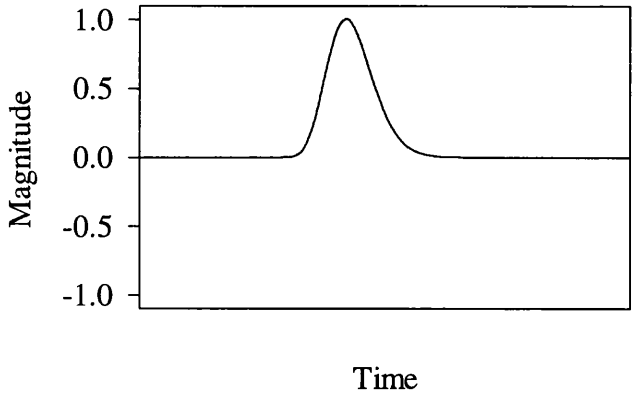
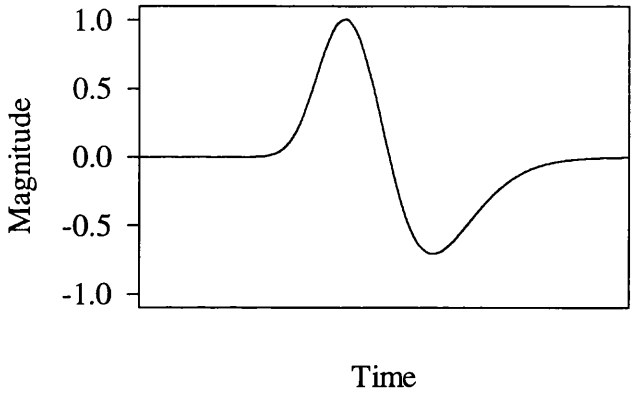


Figure 5.4: Space-time plots of a Γ_L stimulus (a_1), a Γ_C stimulus with a point noise carrier (b_1) and a Φ_C stimulus with a block noise carrier (c_1). The results of "best of both worlds" temporal filtering are shown in a_2, b_2 and c_2. The results of rectification on the filtered images are shown in a_3 , b_3 and c_3.

$f_1(x)$



$f_2'(x)$



$f_1(x) + f_2'(x)$

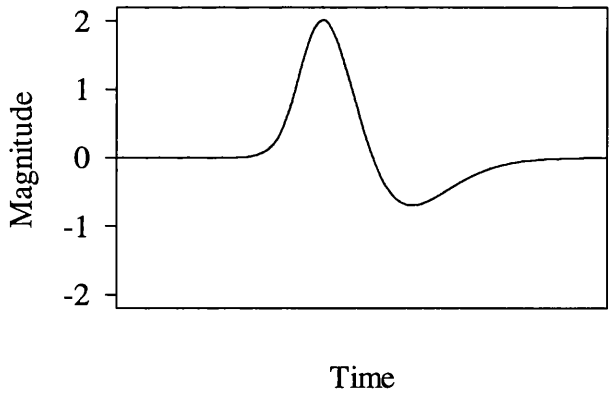


Figure 5.5: A "best of both worlds temporal filter" created from the sum of a log Gaussian ($f_1(x)$) and the differential of a log Gaussian ($f_2'(x)$). The underlying standard deviations are different in order to closely approximate the functions presented by Chubb & Sperling (1989a).

5.3. Results.

The experiments described below investigate perceived direction of motion in luminance and contrast defined polarity reversing bars (Γ_L & Γ_C) and forward stepping bars (Φ_L & Φ_C). Space-time plots of the stimuli are shown in Figure 5.1. The terms *forward* and *reversed* are used to describe perceived motion direction in both luminance and contrast defined stimuli. Forward motion is the direction in which the luminance/contrast bars move by ¼ cycle. Reversed motion refers to a perception of motion in the direction opposite to the displacement. In the descriptions of stimuli presented below, the terms spatial and temporal frequency refer to the frequency of the modulant for both luminance and contrast defined patterns.

5.3.1. Experiment 5.1: Preliminary investigations.

Chubb & Sperling (1989a) explained the percept of forward motion obtained with their Γ_L stimulus by postulating a temporal filtering stage followed by a fullwave rectification-like non-linearity. The initial temporal filter is the average of a lowpass filter and a bandpass filter. Figure 5.4a shows the effect of filtering and rectification on the luminance defined Γ_L stimulus. In the filtering operations shown in Figure 5.4, a filter comprised of the sum of a log Gaussian and the differential of a log Gaussian was employed (see Figure 5.5). In order to generate a filter which has a similar form to that used by Chubb and Sperling (1989a), different time constants for the lowpass and bandpass filters were chosen. It is clear that the filtering and rectification operations allow the extraction of forward motion in the Γ_L stimulus.

Figure 5.4b_1 shows an example of a contrast defined Γ_C stimulus. A two channel model would predict that no coherent motion energy would be detected in the Fourier channel, but that motion detectors in the non-Fourier channel would respond to a filtered and rectified version of the image (Figure 5.4b_3). Since this motion pattern is similar to a Γ_L stimulus, the stimulus should appear to move in the reversed direction. Such a phenomenon may be termed second-order reversed-phi.

In Figure 5.4c the same image processing operations are applied to a Φ modulation of "block noise". The block noise is a 1D dynamic noise that has the same spatial and temporal extents as the bars making up the modulation. It is clear that temporal filtering and subsequent rectification can adequately recover the modulant. If these processes truly underlie human second-order motion processing the subjects should have an unambiguous impression of forward motion in this case.

Eight subjects were tested over two different viewing distances; 1m and 6m. In the following description all terms referring to the spatial dimension are given for the close viewing distance. Γ and Φ patterns were used to modulate dynamic 1D binary "point" noise, which was composed of elements with a width of 0.98 minutes and a temporal extent of 17 msec. In some conditions "point" noise (Figure $5.4b_1$) was replaced with "block" noise (Figure $5.4c_1$). In this case the noise elements had the same temporal and spatial extents as the bars of the modulant. Subjects were also tested on luminance defined Γ_L stimuli. At each distance, 2 bar widths were used, 0.16° and 0.33°, giving spatial frequencies of 1.52 and 0.76 cycles per degree respectively. The temporal frequency of the stimulus was 3.75 Hz, which is the value used by Chubb & Sperling (1989a). The contrast of the luminance defined stimuli was 0.9. The modulation depth for the contrast defined stimuli was 0.8. Subjects were tested with all stimuli at both viewing distances. The order of presentation of classes was random and 30 members of each class were presented in succession. Half of the subjects were tested first at the near viewing distance, then at the far viewing distance. For the remainder the order was reversed.

Subjects made 30 responses to each class of stimulus. The number of responses indicating perceived motion in the forward direction was recorded. A score of 30 means that the subject responded in the forward direction on every presentation of that stimulus, whilst a score of 0 means that the subject responded solely in the reversed direction. A score of 15 indicates no overall bias towards perceiving motion in any direction. The binomial probability distribution gives the probability of *n* or more responses in a particular direction occurring by chance. A score of 21 or more, or 9 or less, can be taken as grounds for rejecting the hypothesis that the subject's performance is due to chance (on the basis of a two tailed test and with 0.05 as the threshold probability). Figure 5.6 shows numbers of subjects reliably seeing motion in a particular direction for luminance defined stimuli and for

1D dynamic point and block noise stimuli. Given the large number of measures, the probability of type I errors occuring is high. Conclusions can only reliably be drawn from the data when a number of subjects respond in the same manner and/or where some definite pattern emerges. The mean responses across subjects for each condition were also compared. The results of this analysis are shown in Table 5.1.

Luminance defined Γ stimuli: With this stimulus, at both viewing distances and both bar widths, the majority of subjects indicated consistent movement in a particular direction. Generally, the direction was not consistent across subjects. This lack of consistency in the perceived direction of motion is reflected in the group data. However, the findings are broadly consonant with those reported by Chubb & Sperling (1989a). At the closer viewing distance there was a greater tendency to see the stimulus moving in the forward direction, and conversely, at the further viewing distance there was a greater tendency to see motion in the reversed direction. As changing the spatial scale is similar to changing the viewing distance, one would expect a similar pattern of results to emerge with changes in scale to that seen with changes in viewing distance. Certainly, most subjects see forward motion when viewing the wide-bar stimulus at the nearer viewing distance and reversed motion when viewing the narrow-bar stimulus at the further viewing distance.

		Lum mod	Point noise		Block Noise	
		Γ_L	Γ_C	Φ_C	Γ_C	Φ_C
Narrow Bars	1 m	15.4	14.9	18.5 *	14.1	17.8 *
	6 m	5.9 †	14.8	15.6	14.0	14.5
Wide Bars	1 m	19.4	11.5 †	21.0 †	13.6	18.4 †
	6 m	8.1	14.0	15.4	15.3	17.4

* Significant deviations from chance calculated by two-tailed *t*-test ($P < 0.05$).

† Significant deviations from chance calculated by two-tailed *t*-test ($P < 0.01$).

Table 5.1: Mean number of forward responses across subjects for Γ_L stimuli and Γ_C and Φ_C stimuli with point and block noise carriers. The maximum number if responses is 30 so the value expected if subjects respond randomly is 15. Numbers greater than 15 show a bias towards forward motion, numbers less than 15 show a bias towards reversed motion.

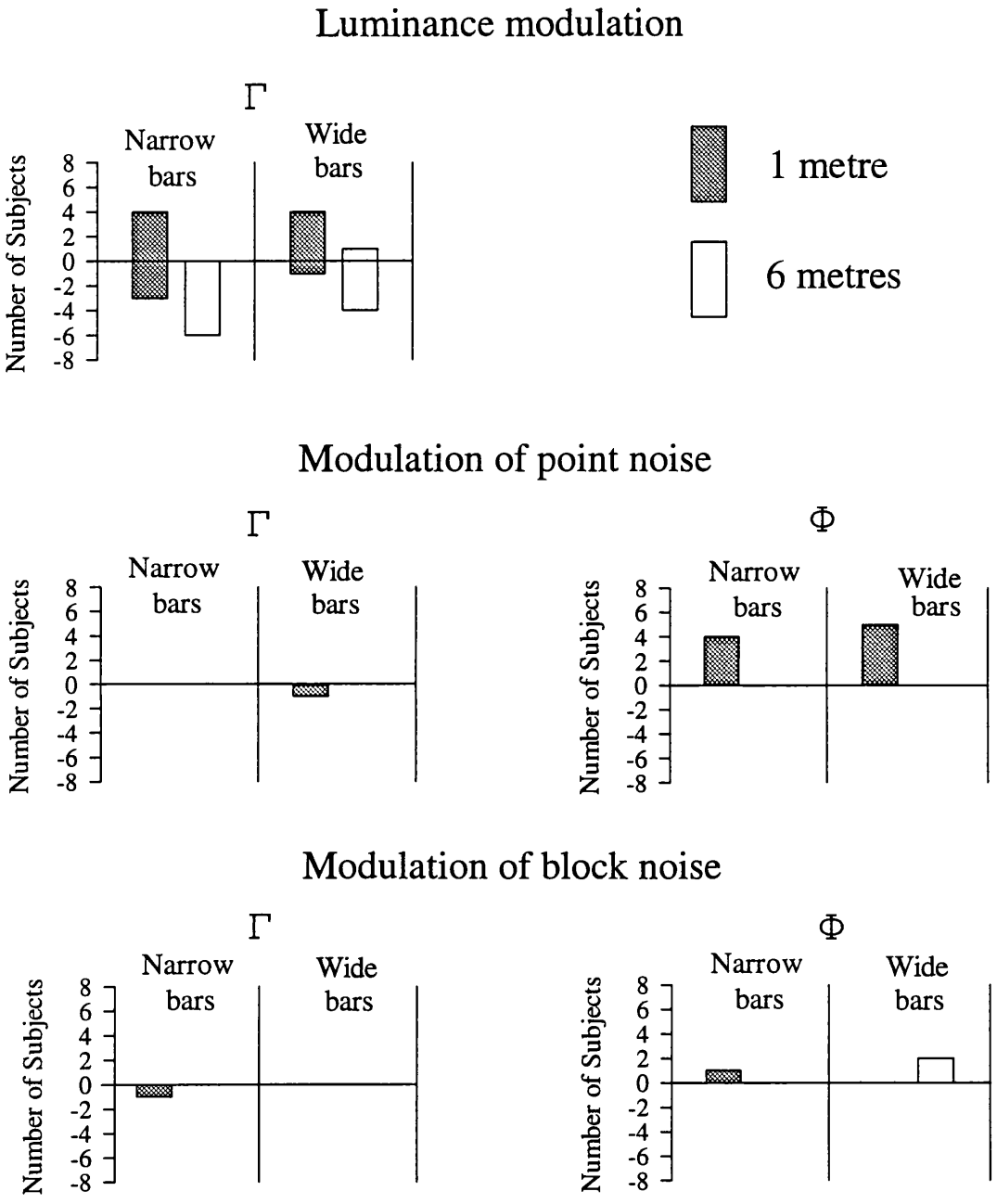


Figure 5.6: Numbers of subjects perceiving forward and reversed motion for luminance and contrast modulations of point and block noise. Forward motion is indicated by bars above the *x-axes*, reversed motion is indicated by those below. The symbols Γ and Φ above the graphs indicate the modulant.

Modulations of 1D dynamic point noise: In general, subjects did not see a consistent direction of motion with Γ modulants. In isolated cases the results of single subjects achieved significance; however, given the number of subjects and the number of stimuli, this is certainly no more than would be expected by chance. The data provide no evidence to support the hypothesis that subjects perceived motion in any consistent direction when viewing these stimuli. This is not the case with Φ modulants. At the closer viewing distance these do elicit a reliable percept of forward motion. Increases in viewing distance substantially reduce this effect but there seems to be no reliable influence of bar width on the perceived direction of motion of the modulant. Results across subjects (Table 5.1) support the conclusions drawn from individual data.

With Γ_C stimuli, there is a small but significant bias in response towards the reversed direction in the group data for the lower spatial frequency at the short viewing distance. In pilot trials, using stimulus durations of 2 seconds, subjects reported tracking the low contrast regions of Γ_C 1D dynamic point noise stimuli. It appears that, if tracking does occur, the probability of jumping from low contrast region to low contrast region in the space-time image may be differentially affected by the contrast of the intervening region. The presence of an intervening medium contrast region is less disruptive than that of an intervening high contrast region, leading to a bias towards the reversed motion percept. Given the small size of the bias, the lack of evidence that these stimuli produce a clear percept of motion in the single subject data, and the high probability of the occurrence of type I errors, it is likely that detection of motion of the Γ_C stimulus reflects some residual eye tracking.

Modulations of 1D dynamic block noise: Results within subjects (Figure 5.6c) show that there is no strong percept of motion in any direction for both Γ_C and Φ_C stimuli. The results across subjects (Table 5.1) show no significant bias for Φ modulants. With the Φ_C stimuli there is some evidence for a small but consistent bias towards seeing motion in the forward direction.

The results for the present experiment can be summarized as follows:

1) The findings of Chubb & Sperling (1989a) are supported.

2) Whilst forward motion is seen in the Φ_C point noise stimuli, this is not the case with Γ_C point noise stimuli, where there is little evidence for any coherent motion percept.
3) In stimuli with 1D dynamic block noise carriers, there is little or no perception of coherent motion in any particular direction.

Whilst the findings of Chubb & Sperling are supported, the effect measured in the current experiment is comparatively weak. The most likely reason for this lies in the differences between the ranges tested. In the present experiment the smallest spatial frequency was 0.76 cycles per degree, and the largest was 9.15 cycles per degree. In Chubb & Sperling's experiment the spatial frequencies ranged from 1.56 cycles per degree to 12.5 cycles per degree. With a higher maximum spatial frequency it is probable that our results would match those of Chubb & Sperling.

The difference between Φ and Γ modulations of 1D dynamic point noise is more problematic. The lack of second-order reversed-phi would seem inconsistent with the findings of Nishida (1993) and indeed of Mather & Murdoch (1996) and Sperling & Lu (1996). However the source of the difference may well lie with the type of carrier used in the present experiment. Smith, Hess & Baker (1994) investigated the effects of carrier on detection and direction identification of modulations of binary noise carriers. The use of a 1D dynamic noise carrier appears to have a seriously detrimental effect on both types of threshold. If motion in Γ_C stimuli was intrinsically more difficult to detect than motion in Φ_C stimuli, then the difference found between these two types of stimulus might simply reflect a difference in threshold. If this is the case then second-order reversed motion may well be elicited by stimuli with carriers other than 1D dynamic noise. Based on the findings of Smith *et al* (1994), static 1D or static 2D noise carriers would be the most likely candidates.

With both Γ_C and Φ_C patterns, the use of a 1D dynamic block noise carrier inhibits the perception of any coherent motion. It is clear from Figure 5.4 that the temporal filtering and rectification operations that successfully recover forward motion in the luminance defined Γ stimulus, can also successfully recover the modulant of a block noise carrier. Therefore the lack of coherent motion in modulations of block noise cannot be explained purely within the context of the proposed second-order channel. However, one may argue that this effect results from the interaction of linear and non-linear channels. For example, the detrimental

effect of the block noise carrier might be explained by proposing that the first-order characteristics of the carrier affect the perception of second-order motion by creating "motion noise" in the first-order channel. This could potentially affect the percept of second-order motion at (or possibly after) an integration stage.

The experiment described above was carried out to replicate data obtained in a slightly different experiment. In this experiment subjects viewed stimuli from 2m and 6m. At the 2m distance the patterns had spatial frequencies of either 1.92 or 0.96 cycles per degree. Apart from the use of an oval window, the procedure and luminance values used were identical to those of the present experiment. From a viewing distance of 1m this elliptical window had the same maximum vertical and horizontal extents as the rectangular window used in experiment 5.1. Again 8 subjects took part in the experiment, 4 of which subsequently took part in the experiment 5.1. Results for the "oval window" experiment are shown in Figure 5.7 and in Table 5.2. The results from the two experiments are very similar.

		Lum mod	Point noise		Block Noise	
		Γ_L	Γ_C	Φ_C	Γ_C	Φ_C
Narrow Bars	2 m	21.3	15.6	16.8	13.6	17.5*
	6 m	10.9	16.4	18.3	14.1	16.3
Wide Bars	2 m	28.4†	13.4*	21.8†	12.5	19.1*
	6 m	18.4	14.5	20.5	14.1	15.9

* Significant deviations from chance calculated by two-tailed *t*-test ($P < 0.05$).

† Significant deviations from chance calculated by two-tailed *t*-test ($P < 0.01$).

Table 5.2: Results from the "oval window" experiment indicating the mean number of forward responses across subjects for Γ_L stimuli and Γ_C and Φ_C stimuli with point and block noise carriers. The maximum number if responses is 30 so the value expected if subjects respond randomly is 15. Numbers greater than 15 show a bias towards forward motion, numbers less than 15 show a bias towards reversed motion.

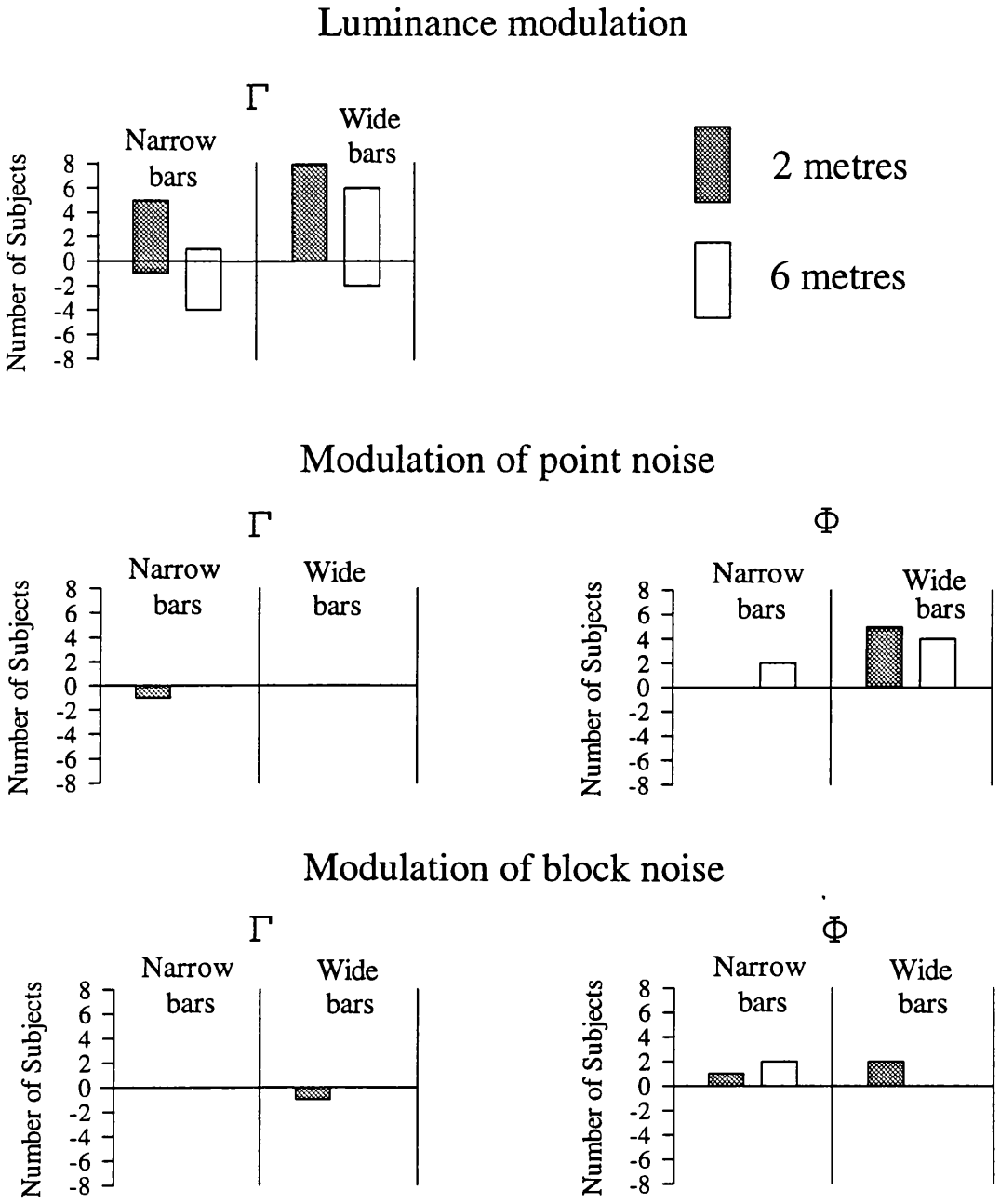


Figure 5.7: Numbers of subjects perceiving forward and reversed motion for luminance and contrast modulations of point and block noise. Forward motion is indicated by bars above the *x-axes*, reversed motion is indicated by those below. The symbols Γ and Φ above the graphs indicate the modulant.

5.3.2. **Experiment 5.2:** Carrier type.

In experiment 5.1, reversed motion in Γ_C stimuli with dynamic 1D carriers was not seen. Forward motion was perceived in Φ_C stimuli. One possible explanation for this difference is that the threshold for detection of motion direction in Γ_C stimuli is higher than that for Φ_C stimuli. If this is the case then the use of a different binary noise carrier may bring the direction of motion in the Γ_C stimulus above threshold.

Eight subjects were tested with Φ_C and Γ_C stimuli over 4 different binary noise carriers: 1D static noise, 1D dynamic noise, 2D static noise and 2D dynamic noise. Figure 5.3 shows sample frames from typical motion sequences to illustrate these stimuli. Luminance defined Φ_L and Γ_L stimuli were also tested. The spatial and temporal frequencies of the modulants in both the luminance and contrast defined stimuli were 0.83 cycles per degree and 3.75 Hz. The width of the noise used in the carrier for the contrast defined stimuli was 3.94 arc minutes. The temporal duration of the dynamic noise elements was 17 msec. Subjects viewed the stimuli from 1m. Unless otherwise specified, the spatial and temporal extents of the noise and the viewing distance should be assumed to take these values for all subsequent experiments described in this chapter. Subjects were presented with 30 trials for each of the 10 stimulus types. The data for all of these were collected in a single session. The stimuli were randomly interleaved; data collection and within subjects analysis were identical to that described for experiment 5.1.

Figure 5.8 shows results in terms of numbers of subjects reporting a coherent direction of motion. In agreement with Nishida (1993), reversed motion is perceived with 2D noise carriers. With Φ_C stimuli, forward motion is readily perceived except for 1D dynamic noise carriers. The threshold hypothesis is clearly supported by these data. The perception of reversed motion in Γ_C stimuli seems particularly sensitive to the dimensionality of the noise with no evidence for any coherent direction of perceived motion for 1D carriers. The results suggest the following ordering in direction discrimination performance for modulations of the various noise carriers:

$$\text{1D dynamic} < \text{1D static} < \text{2D static} < \text{2D dynamic}$$

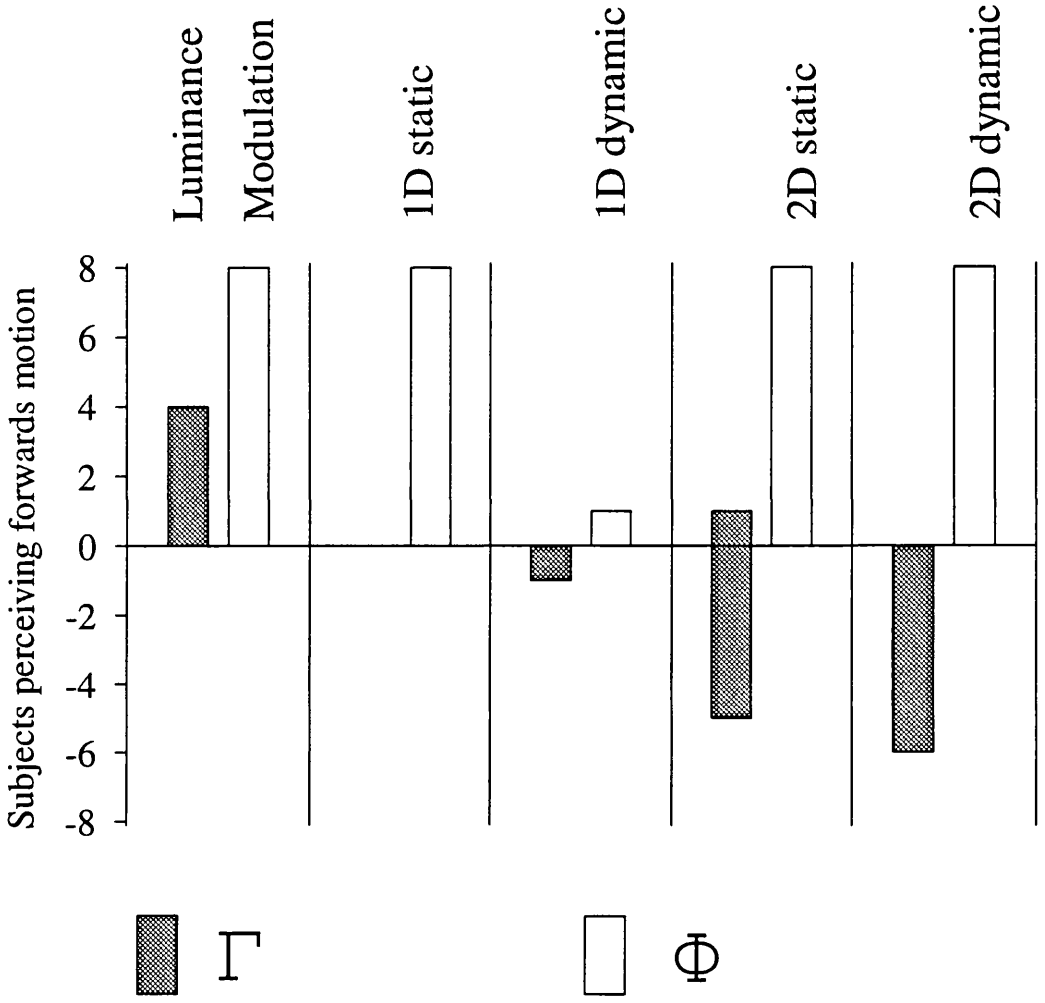


Figure 5.8: Numbers of subjects perceiving forwards motion in luminance- and contrast-defined Γ and Φ stimuli. Negative scores indicate number if subjects perceiving reversed motion. Bar charts identified by noise types show results for contrast-defined stimuli.

In the previous experiment, observers clearly saw motion in Φ_C stimuli with a 1D dynamic carrier; however, in the present chapter there was no reliable perception of motion in this case. The most reasonable explanation for this difference is that, in this experiment, the noise was nearly 4 times wider than that used in experiment 5.1 (3.94 minutes compared to 0.98 minutes).

It is clear that the type of carrier can have a profound effect on the direction of perceived motion. Effects of carrier on the perception of second-order motion have previously been noted (Johnston & Clifford, 1995a) and it would also appear that the nature of the carrier can affect the motion aftereffect for second-order stimuli (Cropper & Hammett, 1996). Additionally, in the spatial domain the orientation of the carrier can affect the perceived orientation of the envelope (McOwan & Johnston, 1996b). It is clear from Figure 5.4b that the temporal filtering and rectification operations described by Chubb and Sperling can adequately recover the modulant. Since the image in Figure 5.4b simply represents one space-time slice through a space-space-time volume, it is also clear that the results of applying the Chubb and Sperling strategy would be the same for 1D and 2D carriers. It is difficult to say why, in the two channel framework, there should be such a profound decrement in performance when 1D carriers are utilised.

5.3.3. **Experiment 5.3:** Modulant temporal frequency.

The results from experiment 5.2 showed that carrier type can have a profound effect on the perception of reversed motion in Γ_C stimuli. Sperling & Lu (1996) showed that, at low temporal frequencies, forward motion could be observed in Γ_C motion sequences. In the present experiment this work was expanded upon by looking at the effect of carrier type on the perception of motion in Γ_C stimuli over a range of modulant temporal frequencies. The effect of temporal frequency on the Γ_L stimulus was also examined. The same 4 binary noise carrier types were used as used in experiment 5.2. Spatial frequency was fixed at 0.24 cycles per degree and 4 levels of temporal frequency were employed: 15 Hz, 7.5 Hz, 3.75 Hz and 1.88 Hz.

Subjects made 50 responses to each level of each stimulus type. The number of responses indicating perceived motion in the forward direction was recorded. A score of 50 means that

the subject responded in the forward direction on every presentation, whilst a score of 0 means that the subject responded solely in the reversed direction. Three subjects, all experienced psychophysical observers, took part in this experiment. Figure 5.9 shows responses to Γ_L and Γ_C stimuli over a range of temporal frequencies for all 3 subjects.

Luminance defined stimuli: All subjects showed a strong reversal. Forward motion was dominant at low temporal frequencies, and reversed motion was dominant at high temporal frequencies. Reversal occurred at about 6 Hz.

Static versus Dynamic carriers: With a static carrier a strong reversal was seen, with forward motion dominant at low temporal frequencies. This reversal appeared to occur at about 7.5 Hz. With a dynamic carrier there was no strong evidence for any percept of forward motion. The percept of reversed motion appears to peak at about 3.75 Hz and, at this frequency, there is a strong difference between the perceived direction of motion for the static and dynamic carriers.

1D versus 2D carriers: Results with 1D carriers appear to be similar to those obtained with 2D carriers, but with responses in the former being a muted version of those in the latter.

The results of Sperling & Lu (1996) are supported, at least in the case of the 2D static carrier. The change from perception of forward motion at low temporal frequencies to reversed motion at high temporal frequencies appears to occur at 7.5 Hz, as was also shown by Sperling & Lu. The use of 1D dynamic noise appears to inhibit any strong motion percept which agrees with the findings in the previous two experiments. In the present experiment, a pronounced effect of carrier type on the perception of motion direction in contrast defined stimuli was obtained. Indeed at 3.75 Hz subjects see motion in the forward direction with a static 2D carrier, and motion in the reversed direction with a dynamic 2D carrier.

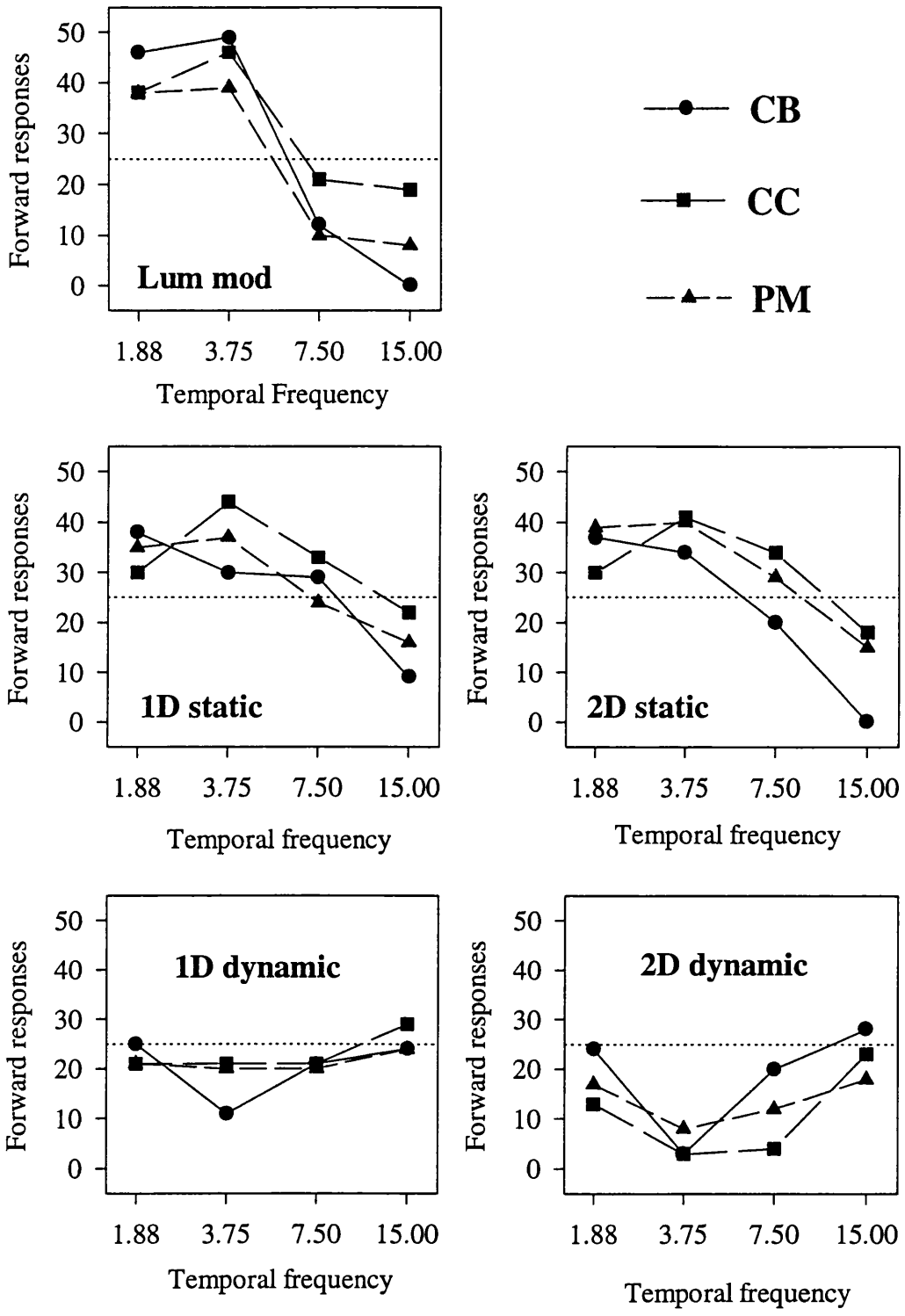


Figure 5.9: Number of forward responses over modulant temporal frequency for Γ_L and Γ_C stimuli. Graphs identified by noise type indicate results for contrast-defined stimuli. Points above the dotted lines show a bias in number of responses towards forward motion, points below show a bias towards reversed motion.

5.3.4. Experiment 5.4: Modulant spatial frequency.

The results from experiment 5.3 showed that a reversal is found over temporal frequency in both luminance and contrast defined Γ stimuli. With Γ_L stimuli, a reversal is also found with increases in viewing distance (Chubb & Sperling, 1989a; see also experiment 5.1). In this experiment, I address the question of whether a similar shift in perceived direction may occur in Γ_C stimuli with a reduction in modulant spatial frequency. This experiment is identical to the previous one except that temporal frequency was fixed at 3.75 Hz and modulant spatial frequency took the following values: 0.12, 0.24, 0.48, 0.95 and 1.91 cycles per degree. Figure 5.10 shows responses to Γ_L and Γ_C motion sequences over spatial frequency for all 3 subjects.

Luminance defined stimuli: There appears to be little effect of spatial frequency. Subjects consistently reported forward motion. Initially, this appears to be at odds with the findings of Chubb & Sperling (1989a) who did obtain reversed motion with this stimulus at long viewing distances. However, the spatial frequency of their stimulus at the far viewing distance (8m) was 12.5 cycles per degree, an order of magnitude greater than the maximum used in the present experiment.

Static versus Dynamic carriers: In modulations of static carriers there is a reversal of perceived motion direction with increasing spatial frequency. Reversed motion is found at the higher spatial frequencies. The reversal appears to occur at about 0.5 cycles per degree. This is directly analogous to the results reported with the luminance defined stimulus by Chubb & Sperling (1989a) except that the reversal occurs at a far lower spatial frequency. With dynamic carriers the percept of forward motion seems to have been completely abolished, at least over the range tested in this experiment. Choice of carrier can therefore completely change the perceived direction of motion of this stimulus.

1D versus 2D carriers: Results with 1D carriers appear to be similar to those obtained with 2D carriers except that the percepts of both forward and reversed motion appear to be strongly inhibited in the case of the former.

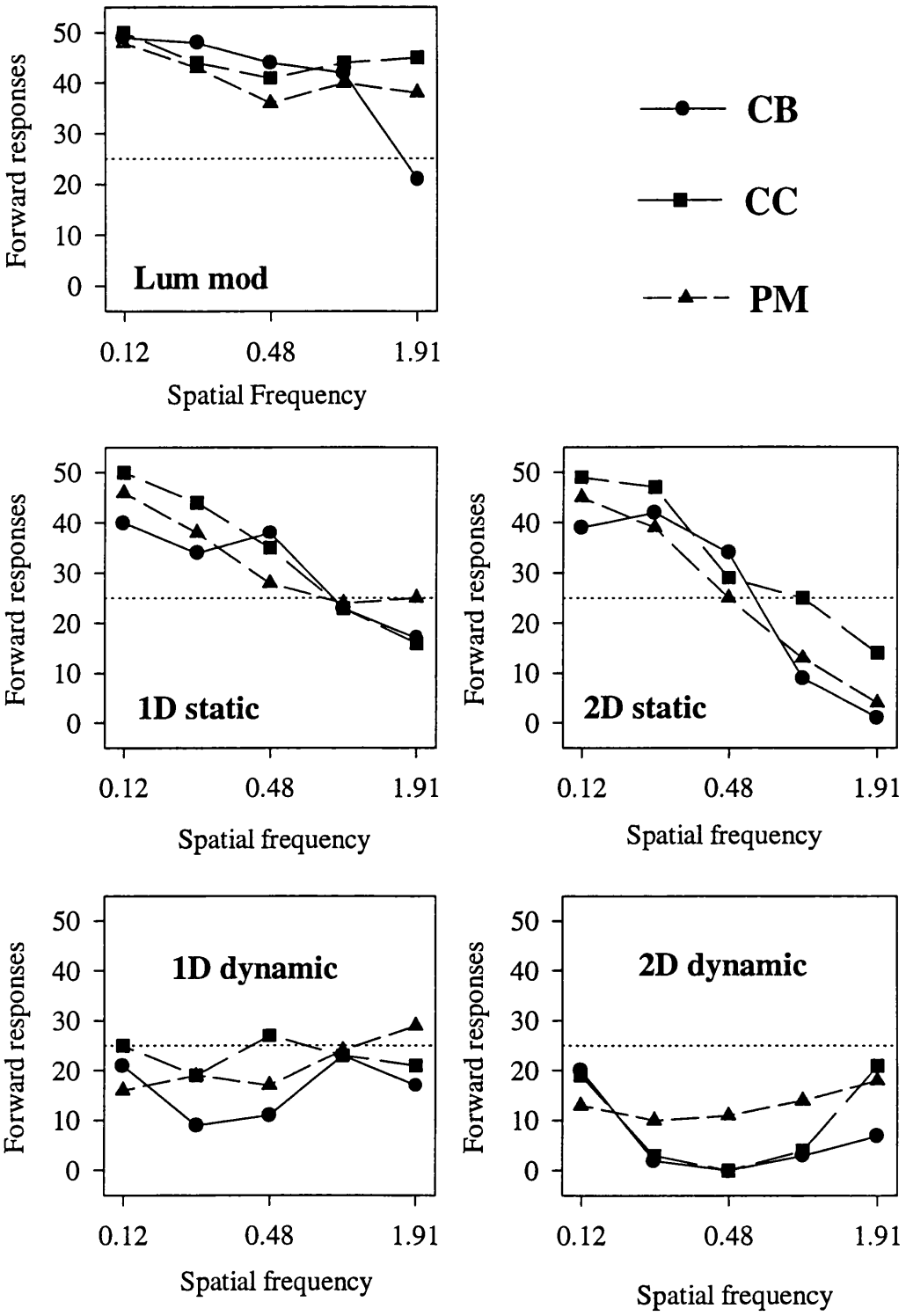


Figure 5.10: Number of forward responses over modulant spatial frequency for Γ_L and Γ_C stimuli. Graphs identified by noise type indicate results for contrast-defined stimuli with that particular carrier.

It appears that, at least in the case of a 2D static carrier, a reversal of motion direction can be obtained with Γ_C stimuli. In the temporal domain the point at which the reversal occurs is similar for both Γ_L and Γ_C stimuli (see experiment 5.3). In the spatial domain, the reversal point is an order of magnitude greater (in terms of spatial frequency) for luminance defined stimuli than for contrast defined stimuli. It would seem therefore that we have a clear quantitative difference between luminance and contrast defined stimuli, with the latter giving rise to changes at a coarser spatial scale. Furthermore, it is certain from these experiments that the nature of the binary noise carrier can have a considerable effect on the perceived direction of motion of the contrast defined stimulus.

In general, a difference in either temporal or spatial frequency tuning is taken as evidence for the existence of two separate mechanisms. For example, Lu & Sperling (1995b) base their identification of a separate third order mechanism principally on the fact that one group of stimuli appears to have a higher temporal frequency cutoff than another group of stimuli. However Smith & Ledgeway (1997b) have recently shown that the temporal frequency cutoff for contrast modulations of static noise is higher than that for contrast modulations of dynamic noise. In fact the pattern of results described by Smith & Ledgeway (1997b) look very similar to those described by Lu & Sperling (1995b) with the modulation of dynamic noise taking the part of Lu & Sperling's "third-order" stimuli. However, modulations of static and dynamic noise are both opaque to standard motion energy analysis (Chubb & Sperling, 1988; see also Chapter 3). Therefore, unless one proposes that the visual system has evolved separate mechanisms for modulations of static and dynamic carriers, one must suppose that the results outlined by Smith & Ledgeway are due to the difference in the carrier type.

If the analysis presented in the paragraph above is correct, then it raises the possibility that the differences in the response over spatial frequency between the luminance and contrast defined motion sequences obtained in this experiment, are due to an effect of the carrier, rather than to some qualitative difference in processing between the two types of stimuli. If this is the case then a single mechanism might show similar dependence on noise whether it is multiplicative, as in the case of contrast defined motion patterns, or additive. Although a contrast defined signal cannot be reproduced simply by adding noise to a luminance modulation, an attempt to examine the effects of a noise carrier may be made by adding

noise with similar characteristics to those of the carrier to a luminance defined stimulus. If the added noise produces similar patterns of results in the first-order stimulus to those seen with the second-order stimulus then it would seem reasonable to assume that the patterns of results is due to the effect of the carrier rather than to some architecturally based difference in processing between first- and second-order stimuli.

5.3.5. Experiment 5.5: Additive noise and temporal frequency.

In this experiment the effects of additive binary noise on the perception of direction of motion in Γ_L stimuli were examined. As the strongest effects in the contrast defined stimuli (see experiments 5.3 & 5.4) are seen with the 2D carriers, only 2D additive noise carriers were employed. In order to allow enough luminance range for the addition of noise, the contrast of the Γ_L stimulus was reduced to 0.19 (in experiment 5.3 the contrast of the Γ_L stimulus was 0.88). Binary noise ($\pm$ 9.73 cd/m^2) was added to the stimulus. The spatial and temporal parameters of the noise were identical to those used in experiments 5.2, 5.3 and 5.4. Modulant spatial frequency was fixed at 0.24 cycles per degree and temporal frequency took the following 4 values: 15 Hz, 7.5 Hz, 3.75 Hz and 1.88 Hz. Figure 5.11 shows the results as a function of temporal frequency for both subjects.

It would appear that the addition of noise also leads to a reversal of perceived motion direction in Γ_L motion patterns. In subject CB, at 3.75 Hz motion was seen in the forward direction with the unadulterated stimulus and in the reversed direction with the addition of 2D dynamic noise. There is no substantial difference between the responses to the luminance defined stimuli presented without added noise found here and those found in experiment 5.3, despite the difference in contrast. The pattern of results for additive noise was very similar to that found with multiplicative noise. In particular, the forward motion elicited by the Γ_L stimulus was eliminated with additive 2D dynamic noise, as was the case for modulation of 2D dynamic noise in experiment 5.3. A similar pattern of results was seen for static noise, although the reduction in the probability of seeing forward motion was less profound.

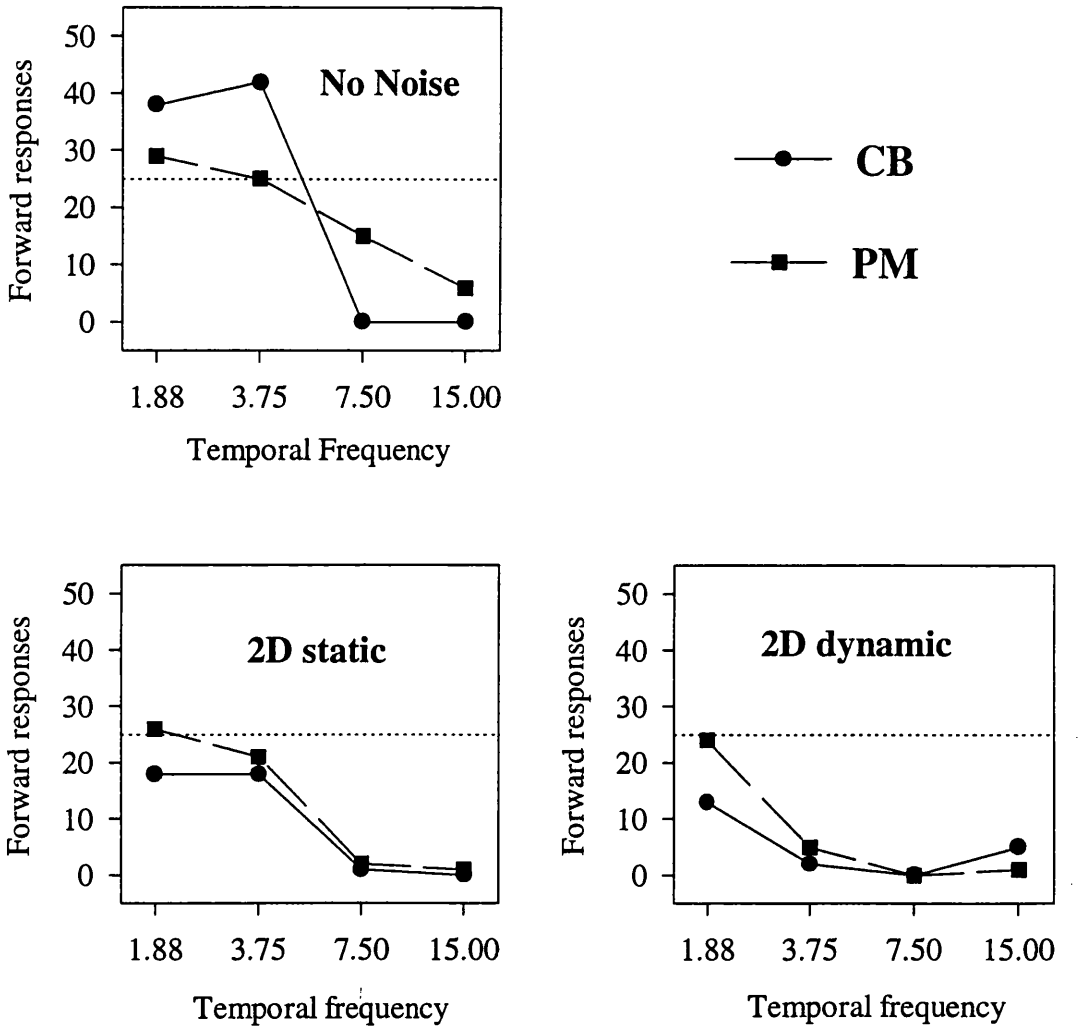


Figure 5.11: Number of forward responses over temporal frequency for Γ_L stimuli with additive noise.

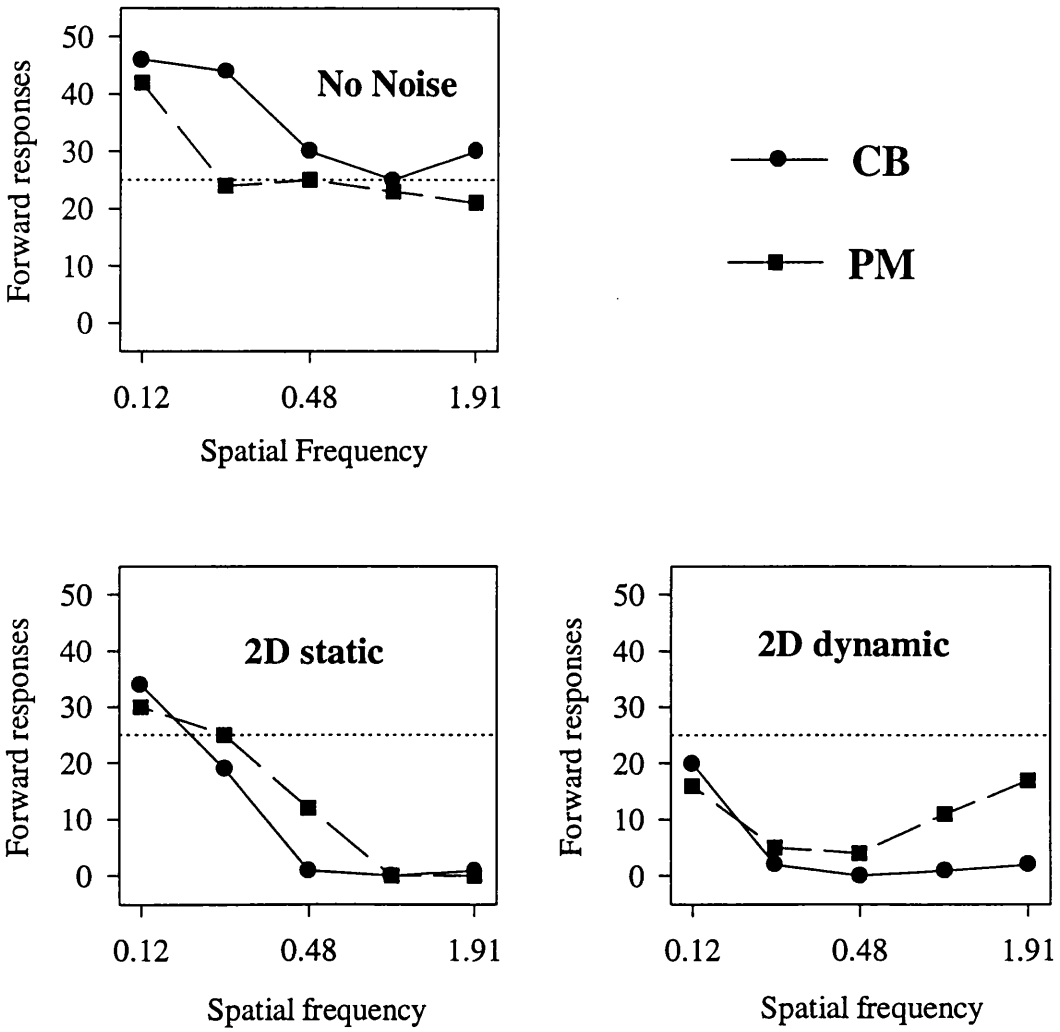


Figure 5.12: Number of forward responses over spatial frequency for Γ_L stimuli with additive noise.

5.3.6. Experiment 5.6: Additive noise and spatial frequency.

The results of the previous experiment are extended by examining the effects of adding static and dynamic 2D noise to Γ_L stimuli as a function of spatial frequency. Temporal frequency was fixed at 3.75 Hz and spatial frequencies of 0.12, 0.24, 0.48, 0.95 and 1.91 cycles per degree were tested. The contrast of the stimuli and amplitude of the additive noise were identical to those in the previous experiment. Data are shown in Figure 5.12.

This experiment confirms the observation made in experiment 5.5 that additive noise can reverse the perceived direction of motion in Γ_L stimuli. Whereas in experiment 5.4 there was no effect of the spatial frequency of the luminance modulation on perceived direction of motion, here we find a strong reduction of forward motion at the higher spatial frequencies in the unadulterated stimulus. Thus there appears to be a strong effect of contrast on perceived direction of motion in stimuli of this type. The differences between additive 2D static and dynamic noise, outlined in experiment 5.5, are far more clearly delineated here. The pattern of results for additive noise was very similar to that found with multiplicative noise for both dynamic and static noise, although there was a weaker tendency to see forward motion overall. In other words, it appears that both dynamic and static noise inhibit the perception of forward motion in this stimulus, but dynamic noise does so to a greater extent. Therefore, it appears that the addition of noise (experiments 5.5 and 5.6) may well give rise to a similar pattern of results to that found for multiplicative noise (experiments 5.3 and 5.4).

5.3.7. Experiment 5.7: Eccentricity.

Sperling & Lu (1996) examined the effect of eccentricity on Γ_C stimuli and found that subjects did not report forward motion for stimuli viewed peripherally (5°). In this experiment, the effect of eccentricity on Γ_C motion sequences with a static 2D carrier at a number of different envelope spatial frequencies was investigated. The effect of eccentricity on Γ_L stimuli was also examined.

Subjects viewed the stimuli from a distance of 60 cm. The rectangle within which the stimuli were displayed had a height of 20.75° and a width of 13.96°. This window was split into

two equal halves from top to bottom; each half was displaced laterally by an equal distance from a central fixation point. The eccentricity was a measure of the distance from the fixation point to the inside edge of each of the two images. When eccentricity is zero there is, therefore, one image in the centre of the screen. When this was the case the fixation spot was overlaid by the stimulus; it was present at all other times. The width of the noise in this experiment was 6.56 minutes. Figure 5.13 shows the results from this experiment and shows the spatial frequencies over which subjects were tested. The temporal frequency of the stimuli was fixed at 3.75 Hz.

In the luminance defined stimulus there was a strong effect of eccentricity, with forward motion perceived in the fovea and reversed motion in the near periphery. This effect is strongly dependent upon spatial frequency. With the contrast defined stimulus there appears to be no such effect. With increasing eccentricity, it appears to be more difficult to see motion in either direction.

Smith *et al* (1994) showed that detection and direction thresholds for first- and second-order stimuli increase at similar rates as eccentricity increases. Their data also suggests that the modulation depth sensitivity curve across spatial frequency shifts downwards as eccentricity is increased, an effect that parallels the change in contrast sensitivity with visual eccentricity (Rovamo & Virsu, 1979; Johnston, 1986, 1987; Drasdo, 1991). Generally, it is assumed that there is an increase in the grain of spatial analysis as one moves outwards from the fovea. If this is the case then increasing eccentricity should be equivalent to increasing the spatial frequency. This analysis is supported by the fact that a reversal of motion direction was found with the luminance defined stimuli over eccentricity. Given this and the fact that Sperling & Lu (1996) did report a reversal of motion direction in contrast defined stimuli with increasing eccentricity, a reversal of perceived direction in the second-order stimuli was expected in this experiment. At present there appears to be no convincing explanation for the absence of such a finding.

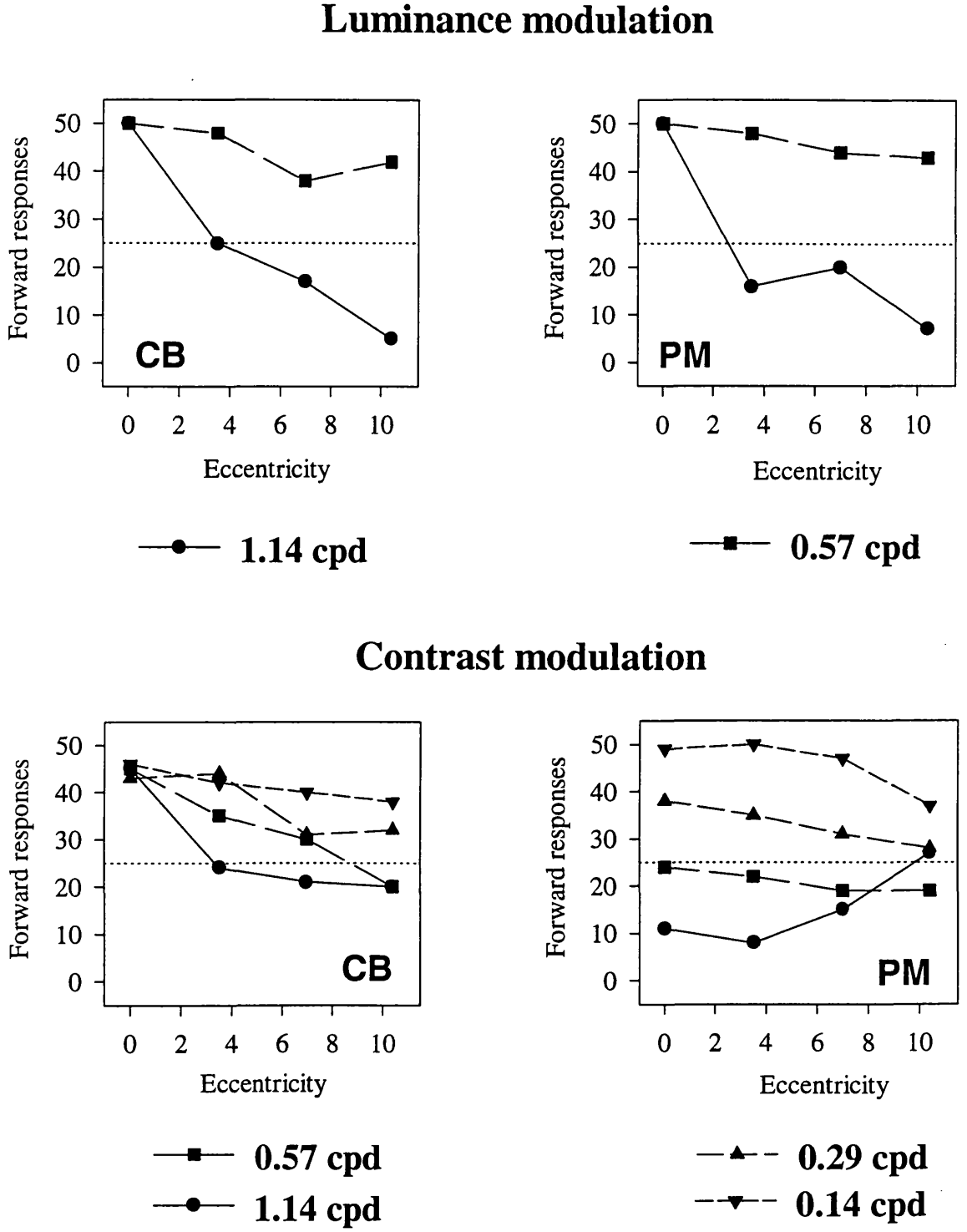


Figure 5.13: Number of forward responses over eccentricity for Γ_L and Γ_C stimuli. Eccentricity is measured in degrees of visual angle.

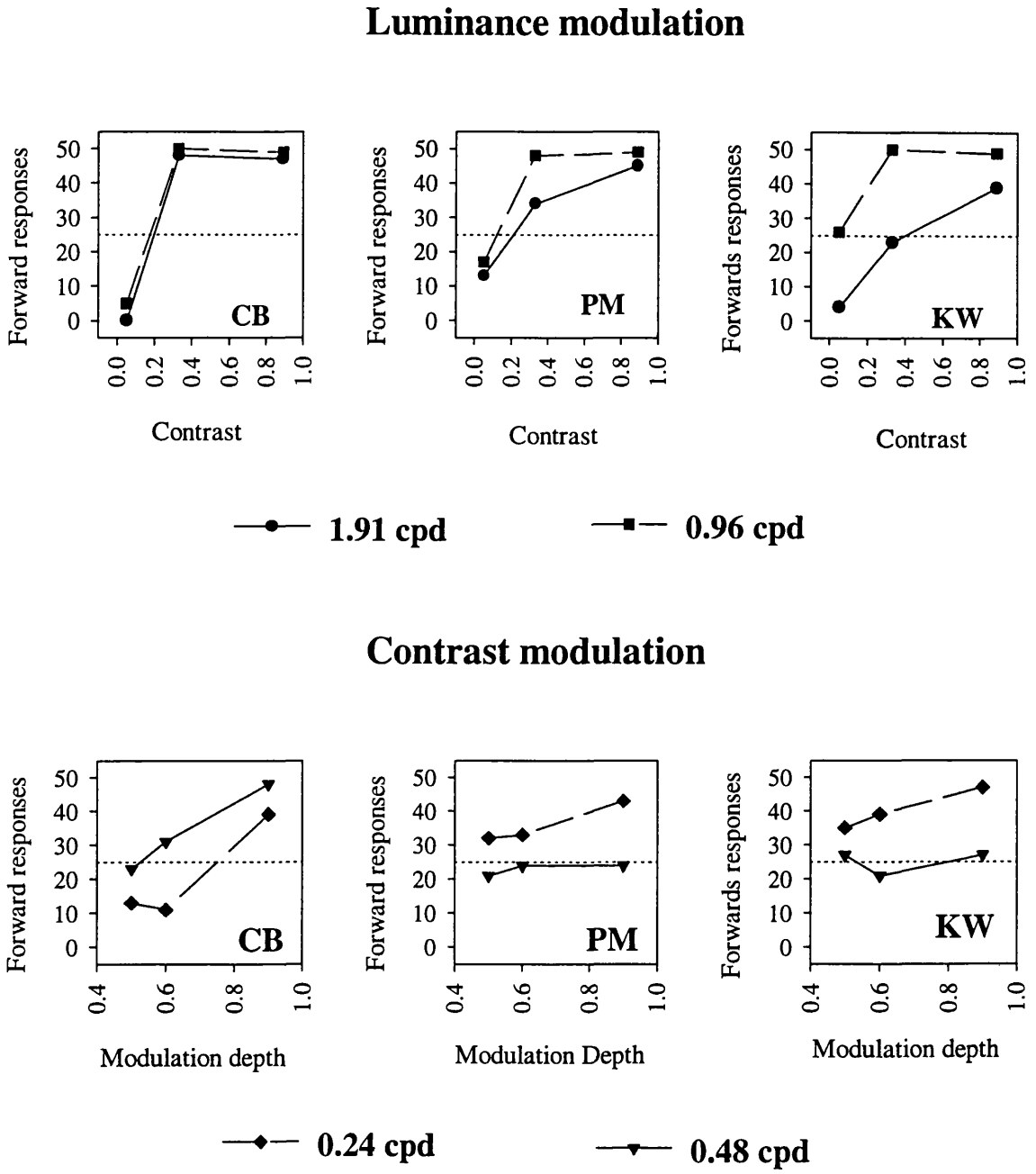


Figure 5.14: Number of forward responses over contrast for Γ_L stimuli and over modulation depth for Γ_C stimuli.

5.3.8. Experiment 5.8: Modulation depth.

In experiment 5.6 it was noted that the percept of forward motion was weaker in low contrast Γ_L stimuli than in high contrast Γ_L stimuli. In this experiment, the effects of contrast on Γ_L motion sequences, and modulation depth on Γ_C stimuli with a 2D static carrier are examined. For the luminance defined stimuli, contrasts of 0.05, 0.33 and 0.89 were used. For the contrast defined stimuli the modulation depths were 0.5, 0.6 and 0.9. Figure 5.14 shows the results for this manipulation and lists the spatial frequencies of the stimuli. Temporal frequency was held at 3.75 Hz.

For the Γ_L stimulus, there is a very strong effect of contrast, with reversed motion perceived at low contrasts and forward motion perceived at high contrasts. Since reversed motion is seen at low contrasts, it is possible that this is an effect of internal noise. If so, it would be analogous to the effects of additive noise found in Experiments 5.5 and 5.6. In the case of the Γ_C stimulus there is a tendency for the perception of forward motion to decline with decreases in modulation depth, but this is not as pronounced as that found with a reduction in contrast of the luminance defined stimulus.

5.4. General Discussion.

Chubb and Sperling (1989a) designed a luminance defined motion sequence, Γ_L, which appears to move forward at close viewing distances but in the opposite direction at greater viewing distances. The results from the experiments described in this chapter show that this reversal can also be generated by increasing spatial frequency, increasing temporal frequency, reducing contrast, adding binary noise and increasing eccentricity. Chubb and Sperling argued that the reversal of apparent motion is evidence for two separate motion systems. They proposed that, in addition to a linear motion energy channel, there exists a second-order or non-Fourier channel which includes a fullwave rectification stage prior to motion analysis. This mechanism is used to explain the percept of forward motion in the Γ_L stimulus whilst reversed motion is explained in terms of the linear channel.

Motion reversal can also be elicited by a contrast defined version of the original Chubb & Sperling (1989a) stimulus, Γ_C. Reversed motion is seen at high spatial frequencies and high

temporal frequencies. The forward motion seen at low spatial and temporal frequencies in this stimulus cannot be accounted for within the two channel framework offered by Chubb & Sperling (1989a). As the stimulus is a contrast modulation of binary noise, it should elicit no consistent response in the first-order channel; however, the filtering and rectification stages in the second-order channel should recover the modulant, or a distorted version of it. As the vast majority of the motion energy in the recovered modulant will indicate motion in the reversed direction, standard motion energy analysis in the second-order channel should signal motion in the reversed direction. Chubb & Sperling (1989a) explained the forward motion in the luminance defined stimulus by proposing a second-order motion mechanism. In a similar vein, Sperling & Lu (1996) proposed an additional "third-order" motion mechanism to account for forward motion seen in the contrast defined stimulus.

It is often thought that, if some motion pattern is not visible to a specific computational mechanism, then another mechanism must exist to account for the perception of that motion. Of course, another approach is also possible. The fact that we can see non-Fourier motion may mean that we have to adjust the existing computational models of low level motion perception rather than add additional mechanisms. In this framework, the percept of non-Fourier motion might be described as epiphenomenal rather than the product of some special system designed to analyse second-order motion. An example of a model in which the perception of second-order motion is epiphenomenal is provided by the Multi-Channel Gradient Model described by Johnston *et al* (1992) and more recently by McOwan & Johnston (1995). Although not designed to recover second-order motion *per se*, this model does successfully describe the perception of motion in contrast modulated sine wave gratings (Johnston & Clifford, 1995a) and can also account for the dependence of perceived motion direction on spatial frequency with the luminance defined Γ_L stimulus (Johnston & Clifford, 1995b). When extended into the motion domain, the spatial primitives approach developed by Watt & Morgan (1985) can also account for this reversal of perceived motion direction over spatial frequency without recourse to an additional mechanism (Morgan, personal communication).

In Watt & Morgan's (1985) model, the halfwave rectified outputs of difference of Gaussian spatial filters of differing scales are combined into two processing streams, an ON stream and an OFF stream. These are then subject to a process in which regions of activity are

identified and labelled as zero bounded regions (ZBRs). The effect of spatial filtering, halfwave rectification and ZBR identification on Chubb & Sperling's (1989a) stimulus are shown in Figures 5.15 and 5.16. When the width of the contrast reversing bars is narrow (Figure 5.15) then the ZBRs appear to step in the direction opposite to that of the displacement. However when the width of the bars is increased (Figure 5.16), we can see that the ZBRs appear to step in the forward direction. It should be noted that the images are only filtered at one spatial scale. However, if the scale of the stimulus is large or small with respect to the initial spatial filters then the results will be qualitatively unchanged.

The purpose of the postulated second-order channel is to extract the modulant in second-order motion sequences. A basic design criterion of any envelope detection device must be the minimisation of the influence of the carrier upon the extraction of envelope motion. The second-order channel is seen as having an initial spatio-temporal filtering stage that precedes the non-linearity. The match between the carrier and these linear filters will determine the subsequent signal strength. It is therefore easy to see how the frequency characteristics of the carrier can affect the amplitude of the recovered modulant. However, the perceived direction of motion in Γ_C patterns shows a profound dependence upon carrier type. Indeed, the choice of carrier can affect the perception of motion in this stimulus to such a degree that forward motion may be seen with a 2D static carrier, whilst only reversed motion is elicited with a 2D dynamic carrier. It is difficult to see how these carrier effects can readily be explained within the context of a second-order channel.

When comparing responses for contrast defined Γ_C stimuli over temporal and spatial frequency, results with 1D carriers appear to be similar to those obtained with 2D carriers, but with responses to the former being a muted version of those to the latter. With 1D carriers, all of the noise has the same spatial orientation as the motion signal. We might therefore expect that 1D carriers should interfere with or mask the motion signal to a greater degree than 2D carriers. If this is correct, then we should find a similar decrement in performance in luminance defined signals with 1D additive noise compared to those with 2D additive noise. At present there does not appear to be any such evidence within the motion domain. However, a study by Rovamo & Kukkonen (1996), examining the detection of static sine wave gratings, showed that detection was more difficult with 1D additive noise than with 2D additive noise.

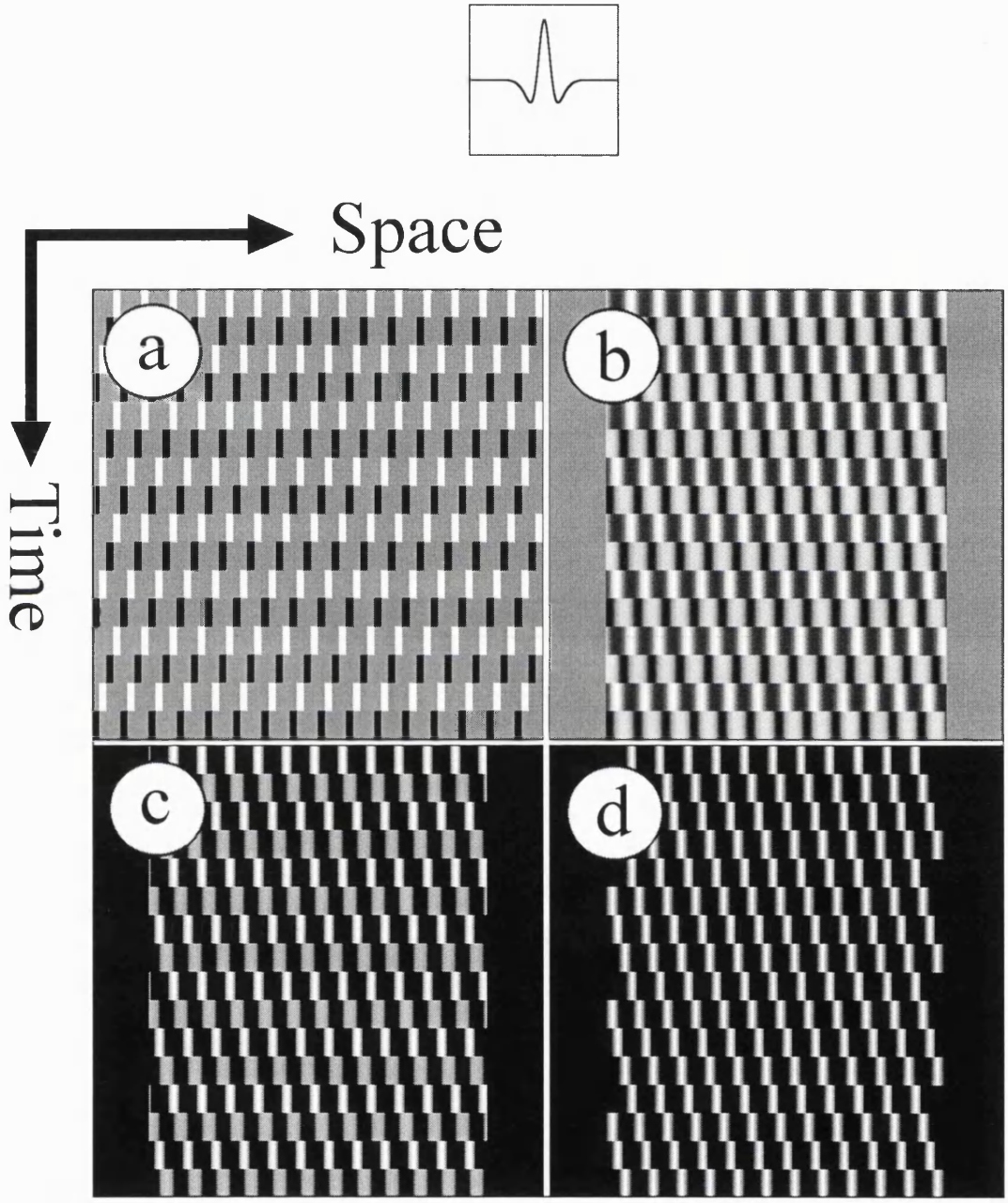


Figure 5.15: (a) Space-time plot showing Chubb & Sperling's (1989a) reversed-phi motion sequence. (b) The result of filtering the stimulus with a difference of Gaussian spatial filter. The profile and spatial scale of the filter are shown above the space-time plots. (c) Halfwave rectification applied to the filtered image. (d) Identification of zero bounded regions in the halfwave rectified image.

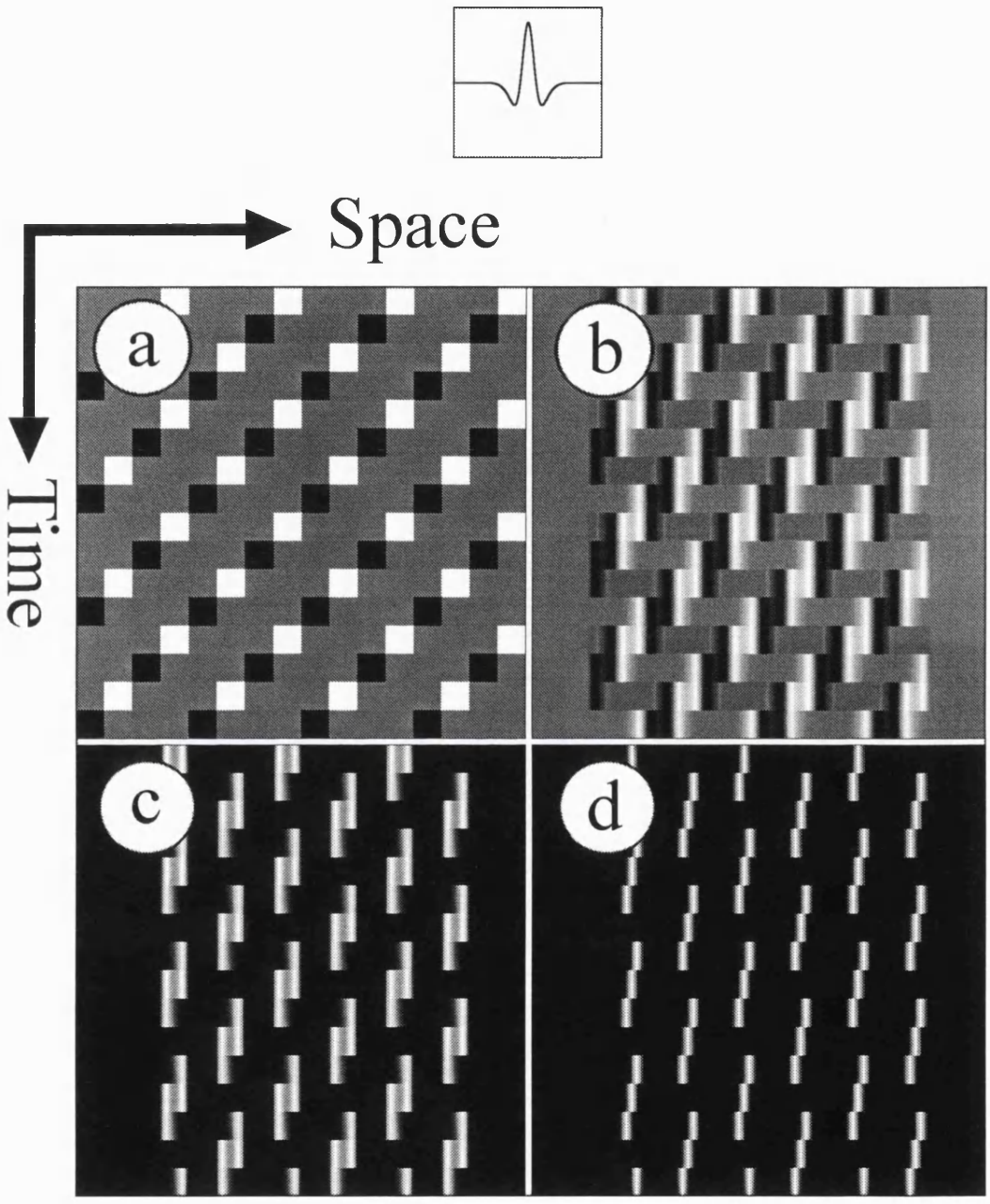


Figure 5.16: (a) Space-time plot showing Chubb & Sperling's (1989a) reversed-phi motion sequence. This is similar to that shown in Figure 5.15a except that the width of the bars has been increased by a factor of 4. (b) The result of filtering the stimulus with a difference of Gaussian spatial filter. The profile and spatial scale of the filter are shown above the space-time plots. (c) Halfwave rectification applied to the filtered image. (d) Identification of zero bounded regions in the halfwave rectified image.

It is clear that the extraction of envelope motion can be radically affected by the nature of the carrier. Such an idea is inimical to models based on demodulative processes, where the whole idea of the proposed mechanism is to extract the modulant whilst effectively discarding the carrier. Such models should predict a degree of separability between carrier and envelope processing. This prediction is not supported by these results. It has been proposed that the computation of velocity may be best understood with reference to the local geometry of the spatio-temporal luminance surface rather than the Fourier composition of the motion pattern (Johnston *et al*, 1992). In the case of a contrast modulation, the local geometry is determined by the properties of both the carrier and the modulant. Of course the form of the carrier will always affect the local surface geometry and therefore the performance of models based upon local geometry will always be susceptible to carrier effects. Indeed, studies that examine "carrier-modulant" separability can potentially discriminate between these two theoretical approaches. For example Cropper & Badcock (1995) showed that the perceived direction of motion in contrast defined plaids was critically dependent upon the orientation bandwidth of the carrier.

The results from experiments detailed in this chapter show that the addition of noise to luminance modulated sequences could radically affect perceived direction of motion. The perceived direction of motion in Γ_L and Γ_C sequences reversed as the spatial frequency of the modulant was increased. The spatial frequency at which reversed motion began to perceptually dominate was estimated to be an order of magnitude greater for high contrast Γ_L sequences than that for Γ_C sequences. This difference was eradicated by reducing the contrast of the Γ_L stimulus and then adding noise. This finding implies that the presence of the noise carrier, rather than some difference in processing, lies at the root of the difference initially observed between the luminance and contrast defined stimuli. This in turn suggests that some of the findings in the literature detailing differences between first- and second-order stimuli may in fact be based on the presence of the carrier rather than some difference in the architecture of motion processing.

The perception of motion in the luminance and contrast defined stimuli appears to follow a similar pattern of response with respect to changes in modulant temporal and spatial frequency. Indeed, the data for contrast defined patterns appear to be similar to that found with noisy low contrast luminance defined patterns. This at least suggests the possibility that

a single mechanism may mediate the perception of both forward and reversed motion, in both the luminance and contrast defined stimuli. A difference in response to two different stimuli should not necessarily be taken as evidence that the two stimuli are processed by different mechanisms. It is possible that two different stimuli may differentially affect a single mechanism.

5.5. Conclusion.

In this chapter the perception of motion direction in luminance and contrast defined Γ stimuli was examined. It was found that both forward and reversed motion could be elicited by both Γ_L and Γ_C stimuli. The forward motion in the latter cannot be explained by traditional two channel models. An additional mechanism is needed to account for the forward motion, necessitating a tripartite architecture. However, it is not clear how the profound carrier dependencies detailed in this chapter could be explained by such an approach. An alternative analysis is offered, in which motion perception in both luminance and contrast defined stimuli is mediated by a single, as yet undetermined, mechanism which shows a similar sensitivity to noise whether that noise is additive or multiplicative.

Chapter Six

Contrast in Motion: evidence for an early non-linearity.

6. Contrast in Motion: evidence for an early non-linearity.

The amplitude of the outputs of linear bandpass filters are independent of the luminance microstructure of noise. This allows a test of the theory of a second-order channel in which images are subject to bandpass filtering prior to rectification. Modulation depth tuning curves for the detection of motion direction in contrast defined motion sequences were compared with tuning curves for stimuli in which the motion was additionally defined by luminance microstructure. In support of the second-order channel hypothesis, no evidence for any effect of luminance microstructure was obtained. No coherent motion was elicited by stimuli that contained motion defined only by luminance microstructure. In additional experiments a more sensitive procedure was developed and an effect of luminance microstructure appeared. It is proposed that the results support the notion of an early non-linearity prior to bandpass filtering.

6.1. Introduction.

In the psychophysics literature, contrast has been defined by a number of metrics (Moulden, Kingdom & Gatley, 1990). The most common of these are the Michelson contrast (C_M) and root mean square contrast (C_{RMS}). C_M is often used to describe the contrast of binary random noise and sine wave gratings whilst C_{RMS} is generally used to describe the contrast of noise. For binary noise the two measures are mathematically equivalent. A number of studies have examined which particular measure might be the more widely applicable. Kukkonen, Rovamo, Tiippana & Näsänen (1993) looked at contrast sensitivity in sine wave gratings and band limited noise. Using C_{RMS} as a metric they found that contrast sensitivity was equal for both of these stimulus types. Tiippana, Näsänen & Rovamo (1994) matched the contrast of a sine wave gratings to supra-threshold complex stimuli consisting of a number of sine wave gratings of the same spatial frequency as the matching stimulus but of different orientations. Again, the two types of stimuli were matched when C_{RMS} was equal but not necessarily when the Michelson contrasts were matched.

Moulden *et al* (1990) studied the utility of several contrast metrics for describing the contrast of random dot images. They measured the effect of an adapting stimulus on the perceived contrast of a test stimulus by getting subjects to match the contrast of a match

stimulus to that of the test stimulus. The adapting stimuli consisted of random noise patterns containing different numbers of grey levels with different distributions whilst the test and match stimuli consisted of binary noise. The study is based upon the assumption that adapting patterns with the same contrasts should have the same effect on the perceived contrast of the test stimulus. Moulden *et al* expressed the contrast of the adapting stimuli using a number of contrast metrics and assumed that the contrast reducing effect on the test stimulus was a direct metric of perceived contrast. Using linear regression, they found that the luminance standard deviation (ie. C_{RMS}) provided the best fit for their data.

So far the evidence merely implies that C_{RMS} may well be a useful general tool for measuring contrast, particularly in random dot patterns. However, as Moulden *et al* tentatively suggest, there may actually be a very good reason why luminance standard deviation should be a good metric for contrast in random dot images. The spectral density of white noise is described by the following equation:

$$N(f_x, f_y, f_t) = C_{RMS}^2 . b_x . b_y . b_t \left(\frac{\sin(\pi f_x . b_x)}{\pi f_x . b_x} \right)^2 \cdot \left(\frac{\sin(\pi f_y . b_y)}{\pi f_y . b_y} \right)^2 \cdot \left(\frac{\sin(\pi f_t . b_t)}{\pi f_t . b_t} \right)^2 \qquad (6.1)$$

where $N(f_x, f_y, f_t)$ is the expected power at coordinates (f_x, f_y, f_t) within frequency space, b_x and b_y are the spatial dimensions of the pixels, and b_t is the temporal extent of the pixels (Legge, Kersten & Burgess, 1987; Kukkonen, Rovamo & Näsänen, 1995). The expected power of any component within the power spectrum is directly proportional to the square of C_{RMS} which means that the amplitude of the component is directly proportional to C_{RMS}. Therefore the expected output of a bandpass filter applied to a noise pattern should be linearly dependent upon the RMS contrast of the noise. Moulden *et al* provide a practical example of this by showing that the convolution product of a Laplacian filter applied to noise patterns is proportional to the luminance standard deviation of the input images. They suggest that given the relationship between luminance standard deviation and the response of bandpass filters, the success of C_{RMS} as a metric is unsurprising.

To reiterate, the expected output of any bandpass filter applied to noise should be proportional to C_{RMS}. Furthermore, if the dimensions of the pixels are identical for two noise patterns, then the expected amplitude of the output from the bandpass filters should be dependent only upon C_{RMS}. Therefore, if two noise patterns have identical root mean square

contrasts but different luminance distributions, the expected output of bandpass filters applied to each of the patterns will be identical.

The standard two channel model consists of a stage of spatially bandpass linear filtering followed by some non-linear process. Generally this non-linearity is seen as consisting of some operation around mean luminance such as fullwave rectification (Lu & Sperling, 1995a, 1996), halfwave rectification (Solomon & Sperling, 1994) or a squaring rectification (Wilson *et al*, 1992). The idea of an operation around mean luminance is particularly appealing because the amplitude of the demodulated envelope is far greater than in the case where some monotonic non-linear function is applied directly to the luminance input. In physiological terms this arrangement is seen as relatively easy to implement. This could be done by simply adding the output of the ON and OFF channels (Sperling, 1989) for fullwave rectification or by acting on the output of either an ON or an OFF channel in the case of halfwave rectification (Watt & Morgan, 1985).

In the following, the term *luminance microstructure* is used to refer to the distribution and number of luminance levels in a stimulus. In this chapter it should be assumed that, when two noise patterns are compared, both the mean luminances of the patterns, and the spatial and temporal dimensions of the pixels within the patterns, are identical. Two patches of noise can differ from one another in both their RMS contrasts and their luminance microstructures. Figure 6.1 shows a schematic diagram of a motion sequence. The sequence contains two types of noise (A and B) which may differ from one another both in contrast and luminance microstructure. This can be conceptualised as a square wave modulation in which both the contrast and microstructure may be modulated. The modulant translates by a quarter of a cycle every frame. The motion in this sequence is therefore defined by contrast and/or luminance microstructure. An example of a motion sequence in which the motion is defined by both of these parameters is shown in Figure 6.2. It can be seen that whilst the low contrast areas consist of binary noise, the high contrast regions contain pixels of four different values (quaternary noise).

In terms of Cavanagh & Mather's (1989) characterisation, motion defined purely by luminance microstructure is defined as second-order. However, microstructure-defined motion should not be readily perceived through the operation of the putative second-order

channel. Because of the initial bandpass spatial filtering stage, contrast defined motion should be dependent only upon the RMS contrast and should be independent of luminance microstructure. If this is the case then there should be no difference in the perception of motion evoked by a modulation of contrast and the perception of motion elicited by a modulation of both contrast and luminance microstructure. However, we can present an alternative hypothesis in which the detection of second-order motion is based upon the effects of some early pre-bandpass filtering non-linearity (ie. the distortion product hypothesis). If this is the case then one might well expect that patterns with different luminance microstructures but the same RMS contrasts would not necessarily be interchangeable.

In this chapter the perception of motion in stimuli defined by changes in contrast and/or luminance microstructure are investigated. Modulation depth tuning curves for the detection of motion direction in contrast defined motion stimuli are compared with those derived from stimuli in which the motion was defined both by contrast and by luminance microstructure. In initial experiments, there was no evidence for any strong effect of luminance microstructure. In stimuli containing microstructure-defined motion that was not accompanied by contrast defined motion, no reliable perception of motion direction was elicited. However, it is shown that these results could be accounted for within the framework of the distortion product hypothesis. In a number of additional experiments a more sensitive procedure was developed. It was shown that luminance microstructure can affect the perception of motion in contrast defined stimuli. Additionally, it was also shown that luminance microstructure can affect perceived contrast of static noise patterns.

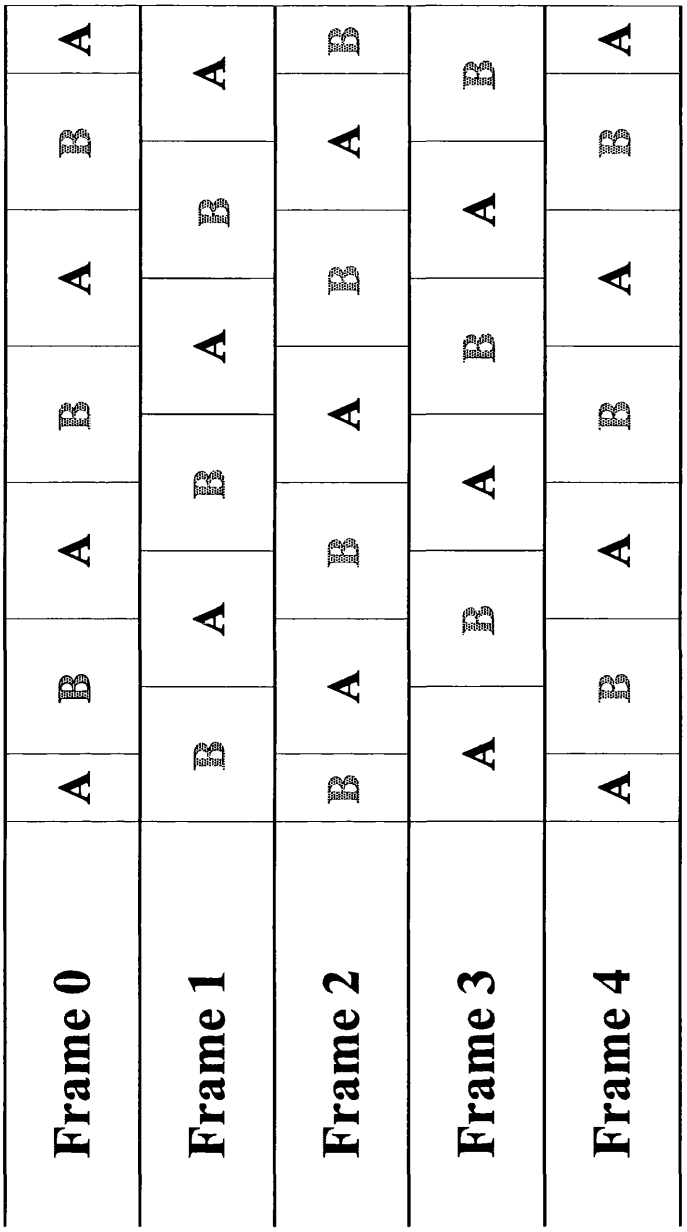


Figure 6.1: Schematic motion sequence showing successive frames containing two noise patterns (A and B) which may differ in contrast and/or luminance microstructure. This may be conceptualised as a modulation of both contrast and luminance microstructure which translates a quarter of a cycle every stimulus frame. One spatial cycle of the modulant contains two adjacent noise patches.

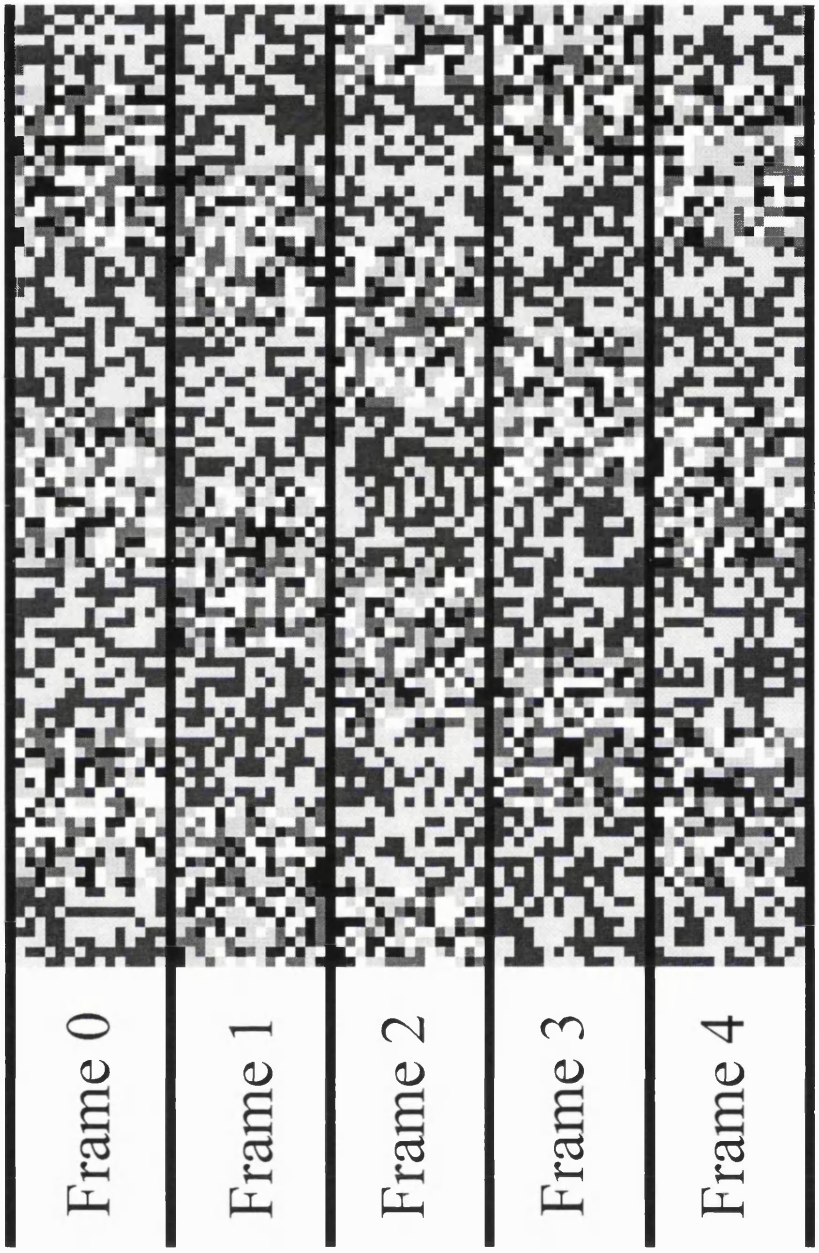


Figure 6.2: Motion sequence with both contrast and luminance microstructure defined motion. High contrast regions contain quaternary noise, low contrast regions contain standard binary noise.

6.2. General Methods.

6.2.1. Stimuli.

The majority of experiments detailed in this study measure modulation depth tuning curves for discriminating the direction of motion. The stimuli consist of two alternating patterns of noise, a target pattern and a comparison pattern. Motion could be defined by both contrast and luminance microstructure. A schematic diagram of this stimulus is shown in Figure 6.1. To take an example, the patches marked with an "A" in Figure 6.1 could contain binary noise with a contrast of 0.5 whilst the patches marked with a "B" could contain quaternary noise with a contrast of 0.7. This is shown in Figure 6.2. These stimuli may be thought of as consisting of a square wave modulant which can modulate both the contrast and the luminance microstructure of noise. One cycle of the spatial cycle of the modulant contains two noise patches. The modulant translates a quarter of a cycle every 67 msec (ie. with a temporal frequency of 3.75 Hz). Each stimulus frame (a period where the modulant does not translate) contains 4 physical frames. Modulation depth tuning curves were measured by holding the contrast of the target pattern constant and varying the contrast of the comparison pattern. For all of the experiments, the comparison pattern was standard binary noise. A number of different noise types were employed for the target pattern. Modulation depth is defined as the following:

$$Modulation\ \ Depth = \frac{C_{target} - C_{comparison}}{C_{target} + C_{comparison}} \tag{6.2}$$

In the present study contrast is always the root mean square contrast and is defined as follows:

$$C_{RMS} = \frac{\sqrt{\sum_{n=1}^{n=h} P(I_n).(I_{mean} - I_n)^2}}{I_{mean}} \tag{6.3}$$

where there are h possible pixel values of intensity I_n each with a probability $P(I_n)$ of occurring and where I_{mean} is the mean luminance. Five types of noise were employed in the present study; standard binary noise, top biased binary noise, bottom balanced binary noise, top biased tertiary noise, bottom balanced tertiary noise and quaternary noise. Examples of

these noise types are shown in Figure 6.3. Each noise type is defined by its composition in terms of the direction and relative displacement from I_{mean} of its luminance components and by the probability of one of those components occurring. This is best illustrated by example. Standard binary contains two luminance components, one at a displacement $+d$ from mean luminance and one at a displacement $-d$ from mean luminance (where $d > 0$). The probability of the $+d$ component occurring is ½ which of course means that the probability of the $-d$ component occurring is also ½. Standard binary noise can therefore be represented by the following list of proportional deviations and associated probabilities:

Standard binary noise: ([+1, ½], [-1, ½]).

Using this terminology the other noise types utilised in the present experiment are defined as follows:

Top biased binary noise: ([+2, ⅓], [-1, ⅔]).

Bottom biased binary noise ([+1, ⅔], [-2, ⅓]).

Top biased tertiary noise: ([+5, ¼], [+1, ¼], [-3, ½]).

Bottom biased tertiary noise: ([+3, ½], [-1, ¼], [-5, ¼]).

Quaternary noise: ([+3, ¼], [+1, ¼], [-1, ¼], [-3, ¼]).

Note that the sum of the products of each pair equals zero. Given the mean luminance, the contrast and the noise type the precise luminance levels can be calculated. In experiments 6.1 and 6.3, dynamic noise patterns were used. The noise pattern in any physical frame was uncorrelated with noise in the previous physical frame. In other experiments, static noise patterns were utilised. In these, the patterns remained static over the four physical frames within each stimulus frame.

When the stimuli contain motion defined by contrast and/or luminance microstructure, the noise pattern on a particular area of the display may be replaced by noise of a different type

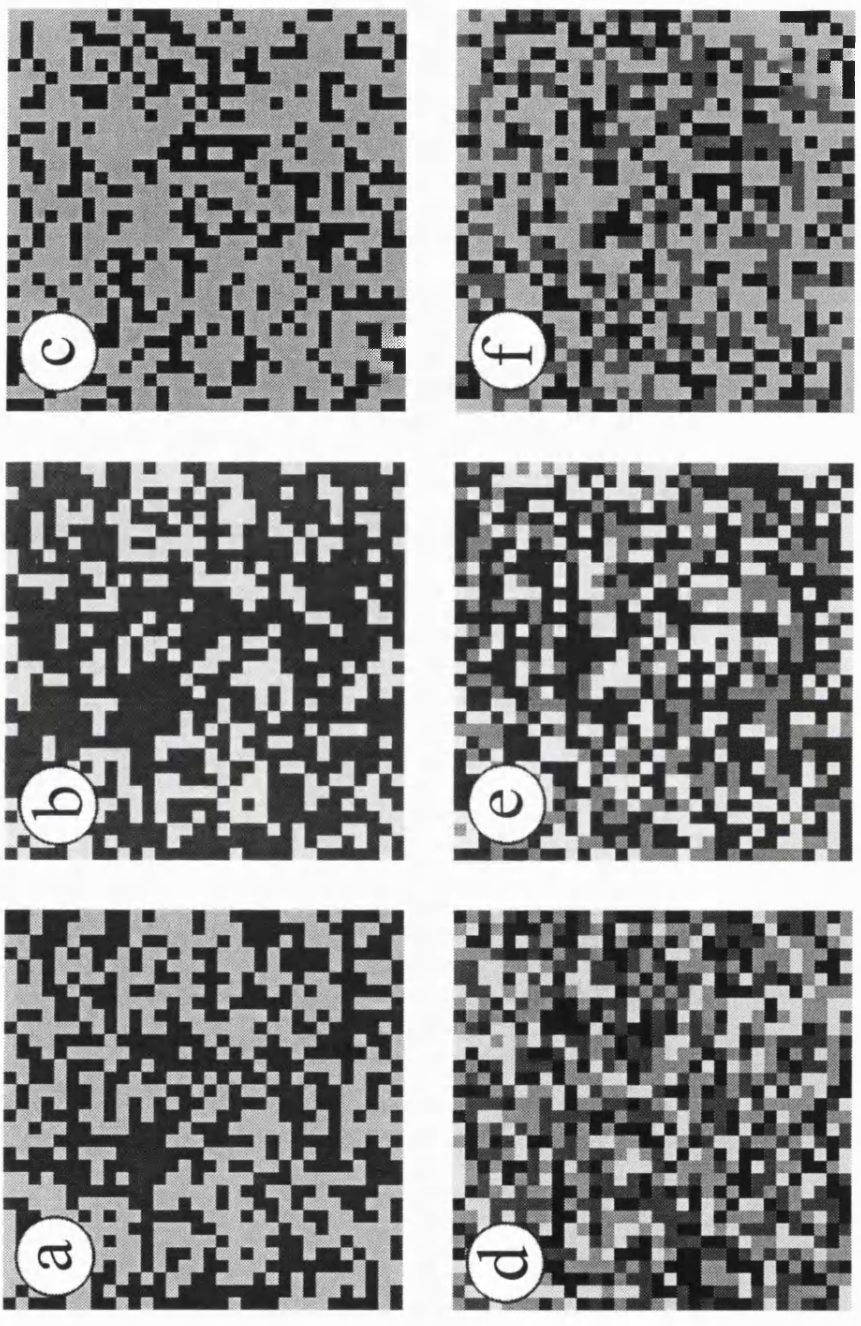


Figure 6.3: Patches of various noise types: (a) standard binary noise, (b) top biased binary noise, (c) bottom biased binary noise, (d) quaternary noise, (e) top biased tertiary noise, and (f) bottom biased tertiary noise.

or contrast. Unless the new noise pattern is exactly the same as the previous one, there will be some degree of flicker. In stimuli which contained static noise patterns, a procedure to reduce the amount of flicker was introduced. For example, let us suppose that a certain area of the screen is covered by standard binary noise but that in the next frame it is replaced by quaternary noise. In this case the underlying pixels should be thought of as consisting of the 4 values 0 to 3 inclusive. When binary noise is mapped over the patch then pixels with values of 0 and 1 become pixels with a luminance of $I_{mean} + x_1$ and pixels with a value of 2 and 3 become pixels with a luminance of $I_{mean} - x_1$. When this patch is replaced by quaternary noise then a value of 0 is mapped to $I_{mean} + 3x_2$, 1 is mapped to $I_{mean} + x_2$, 2 is mapped to $I_{mean} - x_2$ and 3 is mapped to $I_{mean} - 3x_2$. Note that $x_1 > 0$ and $x_2 > 0$, therefore those pixels above mean luminance in one type of noise become pixels above mean luminance in the other type of noise. This reduces the amplitude of the flicker signal when noise of one type replaces that of the other. The mappings between other noise types are similarly arranged and are detailed in Table 6.1. Note that when a stimulus contains static patterns and no microstructure defined motion, then the stimulus is simply a contrast modulation of static noise.

6.2.3. Procedure.

Unless otherwise specified, all experiments contain 2 subjects, CB and KW, both of whom are expert psychophysical observers with normal or corrected to normal vision. All of the experiments described in this chapter were carried out in rooms where the only major source of illumination was the monitor upon which the various stimuli were displayed.

6.2.4. Methodology for Experiments 6.1 to 6.4.

6.2.4.1. Equipment.

Stimuli were displayed on a Manitron monochrome monitor driven by a Matrox IM-1280 graphics card (see Chapter 2 for details). The screen was viewed from a distance of 1 metre, mean luminance was 14.36 cd/m^2. Stimuli were displayed in a rectangular area located in the middle of the screen. From the viewing distance of 1 metre the screen had a width of 20.80° and a height of 16.67°. The height of the patches was equal to the height of the rectangular

	Standard Binary	Standard Binary		Standard Binary	Quaternary
0	$I_{mean} + x_1$	$I_{mean} + x_2$	0	$I_{mean} + x_1$	$I_{mean} + 3x_2$
1	$I_{mean} - x_1$	$I_{mean} - x_2$	1	$I_{mean} + x_1$	$I_{mean} + x_2$
			2	$I_{mean} - x_1$	$I_{mean} - x_2$
			3	$I_{mean} - x_1$	$I_{mean} - 3x_2$
	Standard Binary	Top biased Binary		Standard Binary	Bottom biased Binary
0	$I_{mean} + x_1$	$I_{mean} + 2x_2$	0	$I_{mean} + x_1$	$I_{mean} + x_2$
1	$I_{mean} + x_1$	$I_{mean} + 2x_2$	1	$I_{mean} + x_1$	$I_{mean} + x_2$
2	$I_{mean} + x_1$	$I_{mean} - x_2$	2	$I_{mean} + x_1$	$I_{mean} + x_2$
3	$I_{mean} - x_1$	$I_{mean} - x_2$	3	$I_{mean} - x_1$	$I_{mean} + x_2$
4	$I_{mean} - x_1$	$I_{mean} - x_2$	4	$I_{mean} - x_1$	$I_{mean} - 2x_2$
5	$I_{mean} - x_1$	$I_{mean} - x_2$	5	$I_{mean} - x_1$	$I_{mean} - 2x_2$
	Standard Binary	Top biased Tertiary		Standard Binary	Bottom biased Tertiary
0	$I_{mean} + x_1$	$I_{mean} + 5x_2$	0	$I_{mean} + x_1$	$I_{mean} + 3x_2$
1	$I_{mean} + x_1$	$I_{mean} + x_2$	1	$I_{mean} + x_1$	$I_{mean} + 3x_2$
2	$I_{mean} - x_1$	$I_{mean} - 3x_2$	2	$I_{mean} - x_1$	$I_{mean} - x_2$
3	$I_{mean} - x_1$	$I_{mean} - 3x_2$	3	$I_{mean} - x_1$	$I_{mean} - 5x_2$

Table 6.1: Mappings between luminance values of noise types when they are replaced by one another in stimuli containing static noise patterns. Numbers in columns 1 and 4 refer to the values that the notional underlying noise may take. The luminance values in columns 2 & 3 and 5 & 6 show the luminance values that are mapped over the underlying noise. Note that $x_1 \geq 0$ and $x_2 \geq 0$.

areas within which the stimuli were displayed. The direction of motion of the modulant was horizontal.

In experiments 6.1 and 6.3 the rectangular area was 12.57° wide and 7.86° high. The width of noise patches was 1.58° giving a spatial frequency of 0.32 cycles per degree (cpd) where the length of the cycle is the width of two adjacent noise patches. Noise check size was 2.96 arc minutes. Stimuli were presented for 200 ms which allowed time for two of the quarter cycle steps. The final position of the bars was therefore shifted half a cycle from the start. Subjects could not therefore derive the direction of motion from a knowledge of the start and end positions of the motion sequence.

In experiments 6.2 and 6.4 the width of the height of the rectangular area was increased to 10.46°. The width of the noise patches was 2.10° giving a modulant spatial frequency of 0.24 cpd. Noise check size was 1.97 arc minutes. The stimuli were presented for 600 ms, allowing for 8 quarter cycle steps. As these two experiments utilised static noise, the start and end frames of the sequences were identical.

One potential problem with measuring contrast modulation tuning curves is that the results can easily become contaminated if there is any consistent luminance modulation in the signal. The Matrox IM-1280 graphics card only has 8 bit output, giving a total of 256 discrete luminance values. For a particular luminance level, the usual procedure is to choose the pixel value that gives the closest luminance value to that which is required. Clearly, there may well be a difference between the value required and the value produced. If a number of luminance levels are required to produce a certain contrast then the value of all the luminance levels needed to produce that contrast may vary by some small amount. The actual expected mean luminance level of the noise pattern may therefore differ from the required expected mean luminance level of the noise pattern.

The motion stimuli in this chapter consist of alternating noise patterns (see Figures 6.1 and 6.2). If there is a consistent difference between the actual expected mean luminance levels of the target and comparison patterns, then luminance defined motion will be introduced into the signal. To minimise the possibility of this occurring, the difference between the actual expected mean levels of the two patterns needed to be minimised. This was achieved using

a simple nearest neighbours search procedure (in terms of the luminance values that could be produced) so that the actual expected mean luminance of the comparison pattern would be close to the actual expected mean luminance of the target pattern. The use of this procedure meant that there was some jitter in the comparison contrast which carries through into the modulation depths. Therefore in the results detailed below, the modulation depths plotted are those calculated from the actual discrete values used in the stimuli.

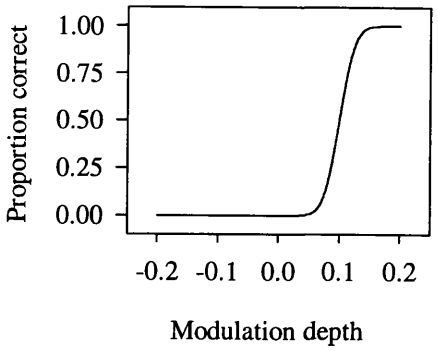
6.2.4.2. Procedures.

Experiments 6.1 to 6.4 each consist of one or more sets of trials. Each set of trials contains two runs of a particular stimulus type where a stimulus type is defined by the two noise types that make up the stimuli. A run consists of a single stimulus type at a number of predetermined levels, a level being defined by the modulation depth and the contrast of the target pattern. Within a set, the target patterns of the two runs have the same contrasts as one another. The members of each of the sets are repeated a number of times so that the time taken to run a whole set takes about 20 minutes. Each set was repeated until 100 trials per level per stimulus type had been completed. The trials within a set were stacked and chosen at random from that stack. For each trial the stimulus was generated online. Starting position and direction of motion were randomized. Subjects were asked to fixate in the centre of the screen and judge whether the stimuli were moving to the right or to the left. Subjects responded by pressing the left/right arrow keys of a PC keyboard. Feedback was provided in all four experiments for both correct and incorrect responses. There was a minimum gap of 1 second between trials.

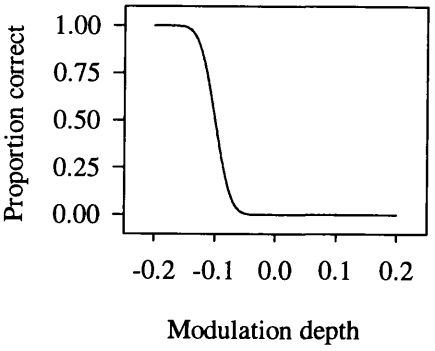
6.2.4.3. Data collection and analysis.

Data consisted of a number of scores out of 100 which indicate the number of correct responses for a particular level on a particular stimulus type. From this the proportion of correct responses can be calculated. In the definition of modulation depth given in equation 6.2, modulation depth is negative when the comparison contrast is greater than the target contrast and positive when the comparison contrast is less than the target contrast. As the absolute value of modulation depth increases, so too should the number of correct responses made by the subjects. On a graph showing proportion of correct responses over modulation

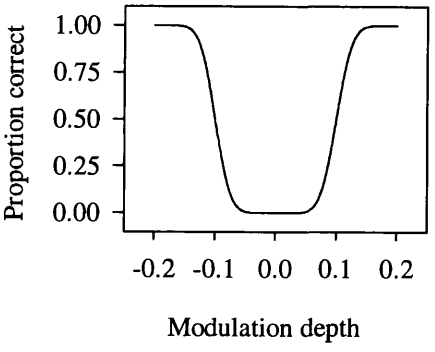
(a) $\text{err}_1(x)$



(b) $1 - \text{err}_2(x)$



(c) $1 + \text{err}_1(x) - \text{err}_2(x)$



(d) $\dfrac{2 + \text{err}_1(x) - \text{err}_2(x)}{2}$

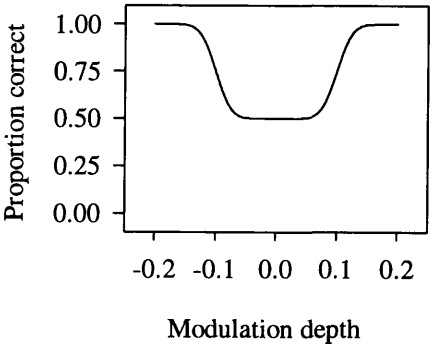


Figure 6.4: The functions $\text{err}_1(x)$ and $\text{err}_2(x)$ are approximations to the integral of a Gaussian with means of μ_1 and μ_2 and standard deviations of σ_1 and σ_2 respectively. The steps (a) to (d) show how the two psychometric functions may be combined to create a U-shaped curve with a minimum of 0.5 and a maximum of 1.0. In the example μ_1 equals 0.1, μ_2 equals -0.1 and both σ_1 and σ_2 equal 0.02.

depth one would therefore expect to find a U-shaped curve. Traditionally each side of the curve would be fitted by a psychometric function. Therefore the following curve was fitted to the data for each stimulus type:

$$f(x) = \frac{2 + err_1(x) - err_2(x)}{2}, \qquad 6.4$$

where $err_1(x)$ and $err_2(x)$ are approximations to an integral of a Gaussian with means of μ_1 and μ_2 and standard deviations of σ_1 and σ_2 respectively. Figure 6.4 shows how the function shown above is formed from two integral of Gaussian functions, and may describe a U-shaped curve. The first regression was followed by a number of iterations of weight generation and weighted regression. Weights were generated in the same manner as those generated for probit analysis (Finney, 1971). The process was halted when further iterations appeared to have little or no effect.

6.3. Results and Discussion.

The following two experiments measure and compare modulation depth tuning curves for the discrimination of motion direction in contrast defined motion stimuli with and without motion defined by luminance microstructure. If second-order motion is perceived through the operation of an initial bandpass filtering operation, then the presence of microstructure-defined motion should have little or no effect. Modulation depth tuning curves for the two types of stimulus should appear identical. However, if second-order motion is perceived through the operation of an early non-linearity, then the mean levels of two different types of noise are not necessarily identical once their component luminance values have been transformed through some non-linear process. If this is the case then we might well expect that modulation depth tuning curves for stimuli containing microstructure-defined motion are shifted with respect to the modulation depth tuning curves for stimuli that contain only contrast defined motion.

6.3.1.1. Experiment 6.1: Modulation depth tuning curves - dynamic patterns.

This experiment contained two sets of trials. In the first set, both runs had target contrasts of 0.6 and in the second set both runs had target contrasts of 0.4. All stimuli consisted of

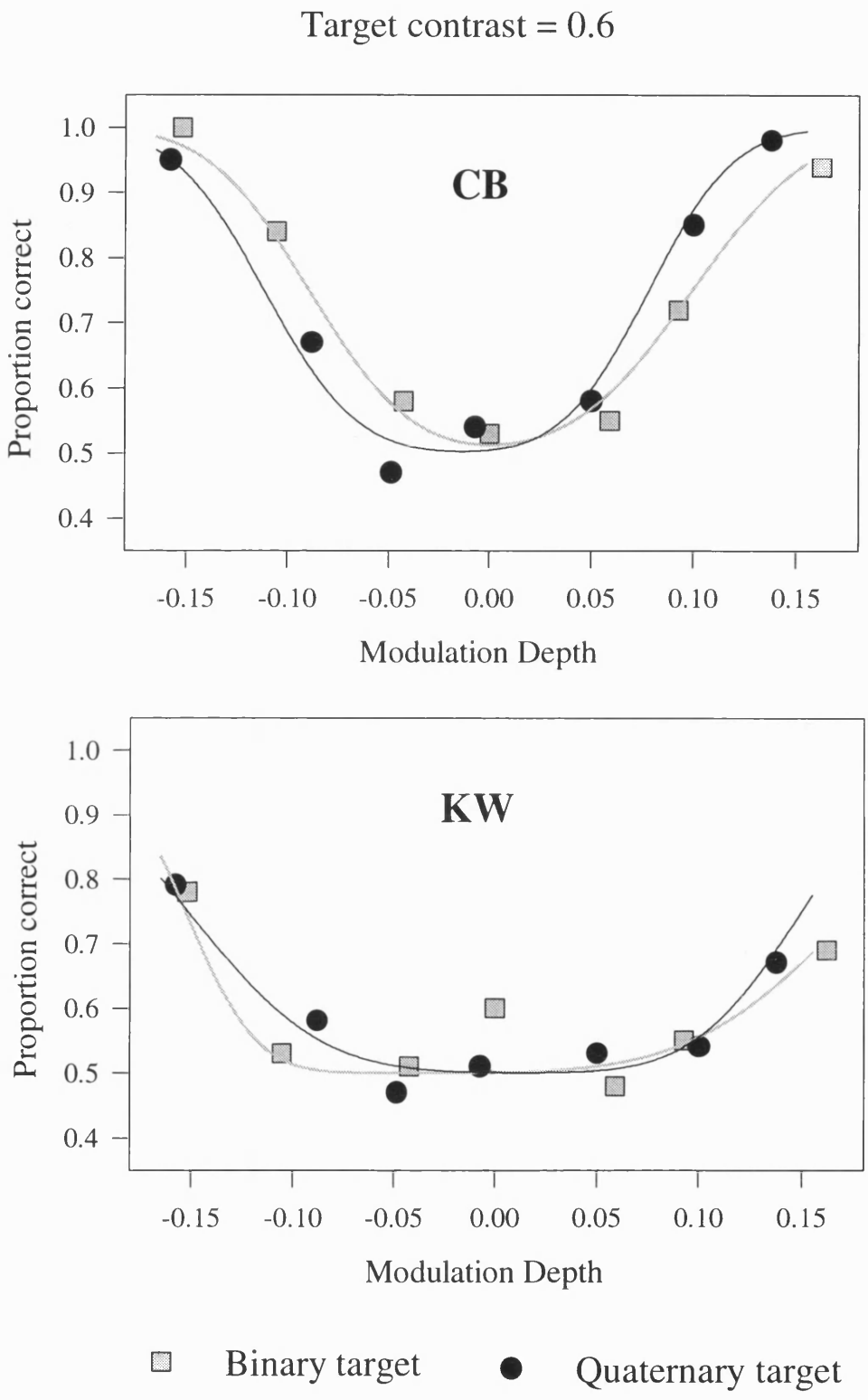


Figure 6.5: Results from experiment 6.1 showing modulation depth tuning curves for the detection of motion direction. The grey line shows the result of fitting equation 6.4 to the data derived from stimuli with standard binary targets. The black line shows the result of fitting the curve to the data derived from stimuli with quaternary targets.

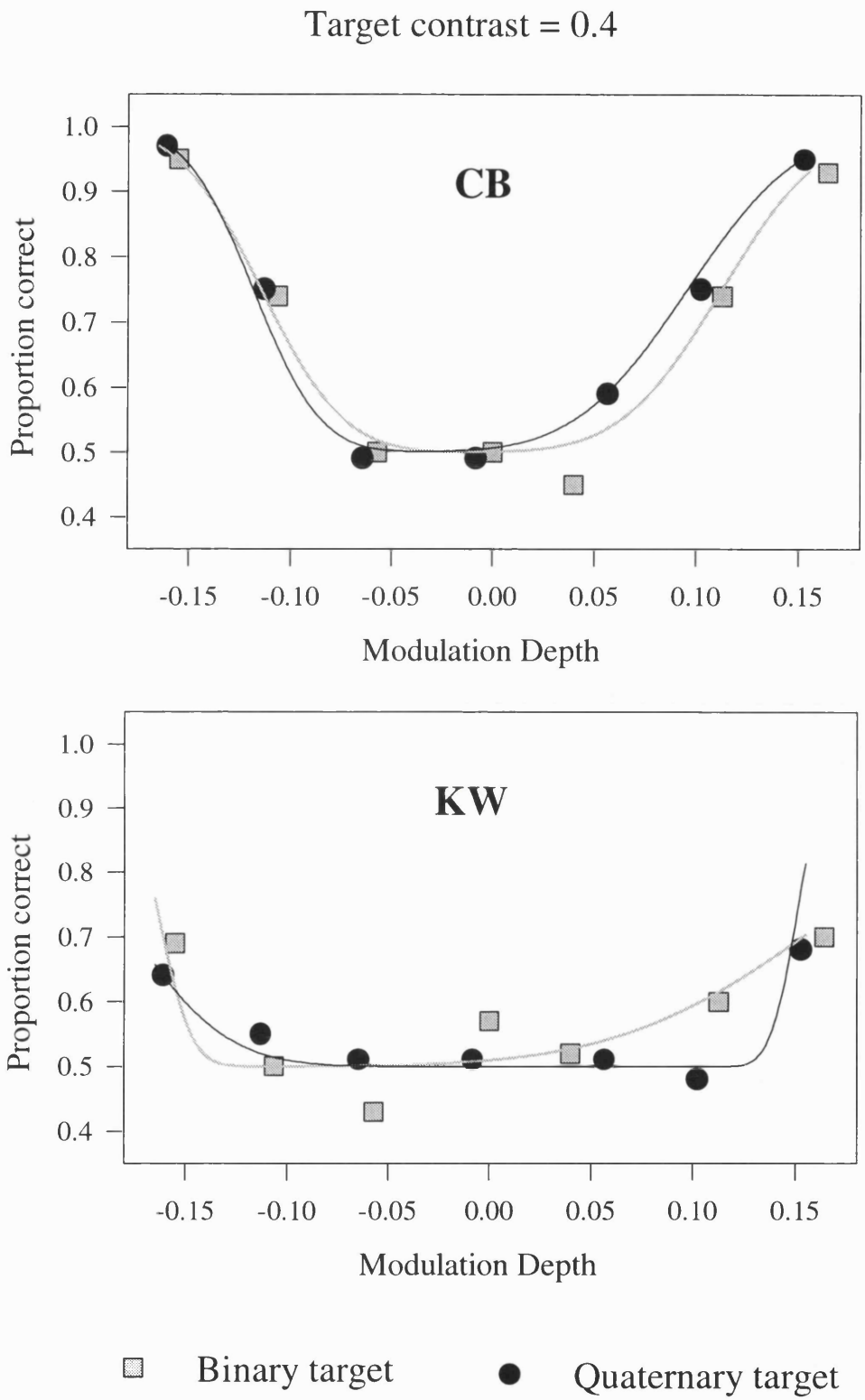


Figure 6.6: Results from experiment 6.1 showing modulation depth tuning curves for the detection of motion direction. The grey line shows the result of fitting equation 6.4 to the data derived from stimuli with standard binary targets. The black line shows the result of fitting the curve to the data derived from stimuli with quaternary targets.

dynamic patterns. Each set contained two runs, one with a standard binary noise target and one with a quaternary noise target. Each run contained 7 levels of modulation depth. These data are plotted in Figures 6.5 and 6.6. For both the 0.6 and 0.4 conditions the curves for subject CB show a slight rightwards shift when a binary noise target is replaced by quaternary noise. When modulation depth is positive, the comparison contrast is less that the target contrast. This implies that the quaternary noise has slightly less apparent contrast than the binary noise even though their RMS contrasts are equal. This effect is very small and is certainly not backed up by the results from subject KW. With the latter the results are too noisy to draw any strong conclusions. There also appears to be an effect of target contrast with, in both subjects, the curves being narrower for the 0.6 condition than the 0.4 condition.

6.3.1.2. Experiment 6.2: Modulation depth tuning curves - static patterns.

In the previous experiment it was clear that subject KW had a great deal of difficulty in accurately completing the task. This experiment is essentially a copy of experiment 6.2 but done with a static noise patterns and with a longer stimulus presentation time. The results from this experiment are shown in Figures 6.7 and 6.8. It is clear that the levels of performance from both subjects are far better than those obtained in experiment 6.1. The slight rightward shift with the quaternary noise target is not present in these data although both subjects do show the narrowing of the U-shape in the 0.6 condition compared to the 0.4 condition. It would appear that there are no systematic differences between the data for quaternary noise targets and the data for standard binary noise targets.

6.3.1.3. Discussion of results from Experiments 6.1 & 6.2.

No effect of luminance microstructure was obtained in experiments 6.1 and 6.2. It appears that when motion is defined by microstructure, but is not defined by differences in contrast, no consistent motion percept is elicited. Luminance microstructure defined motion therefore appears to be one type of second-order motion that we cannot readily perceive. When static, however, the different patterns are readily distinguishable from one another, even when their contrasts are equal. Clearly, at least with the stimulus parameters employed in these experiments, there is no evidence for the involvement of an active attentional feature

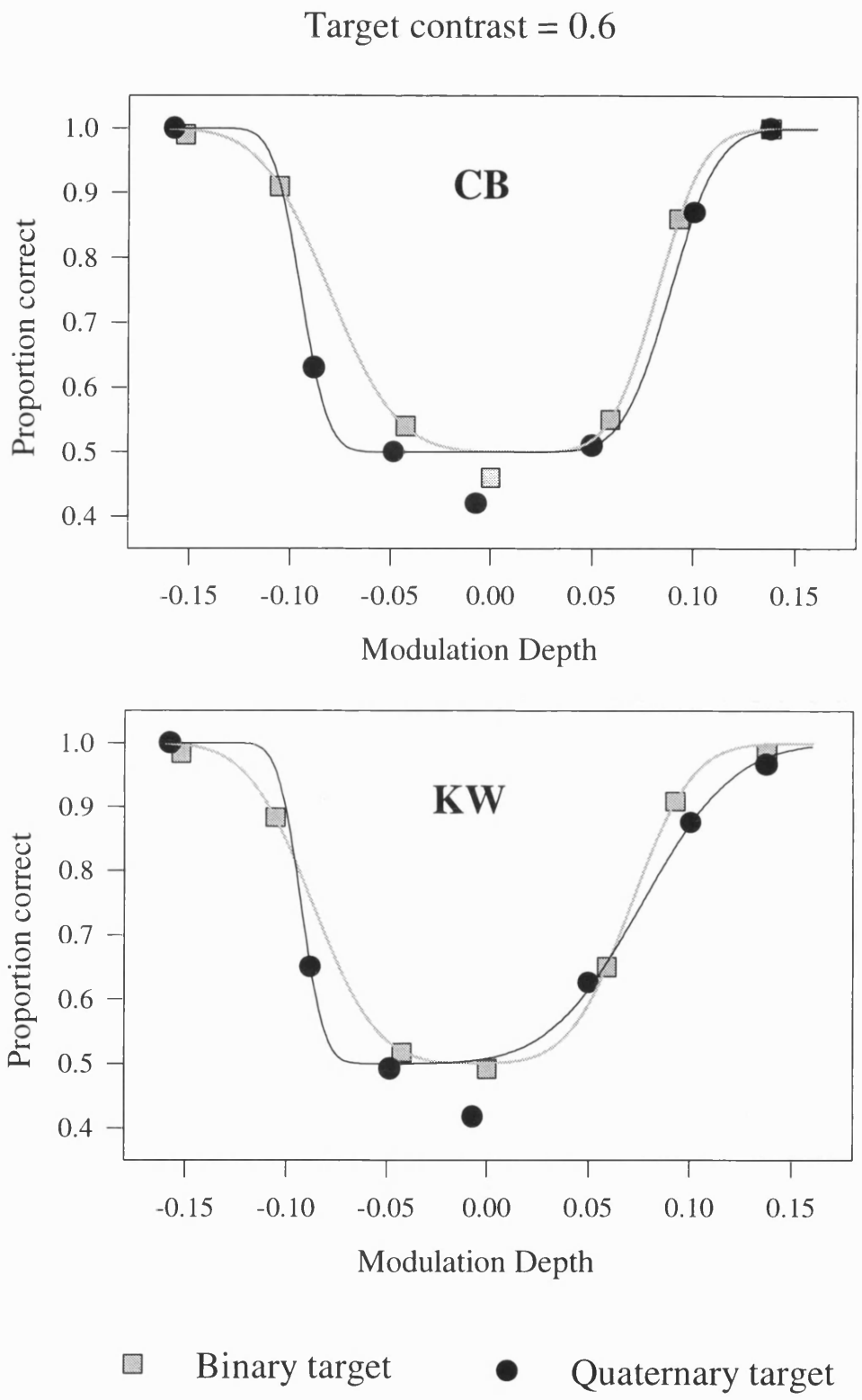


Figure 6.7: Results from experiment 6.2 showing modulation depth tuning curves for the detection of motion direction. The grey line shows the result of fitting equation 6.4 to the data derived from stimuli with standard binary targets. The black line shows the result of fitting the curve to the data derived from stimuli with quaternary targets.

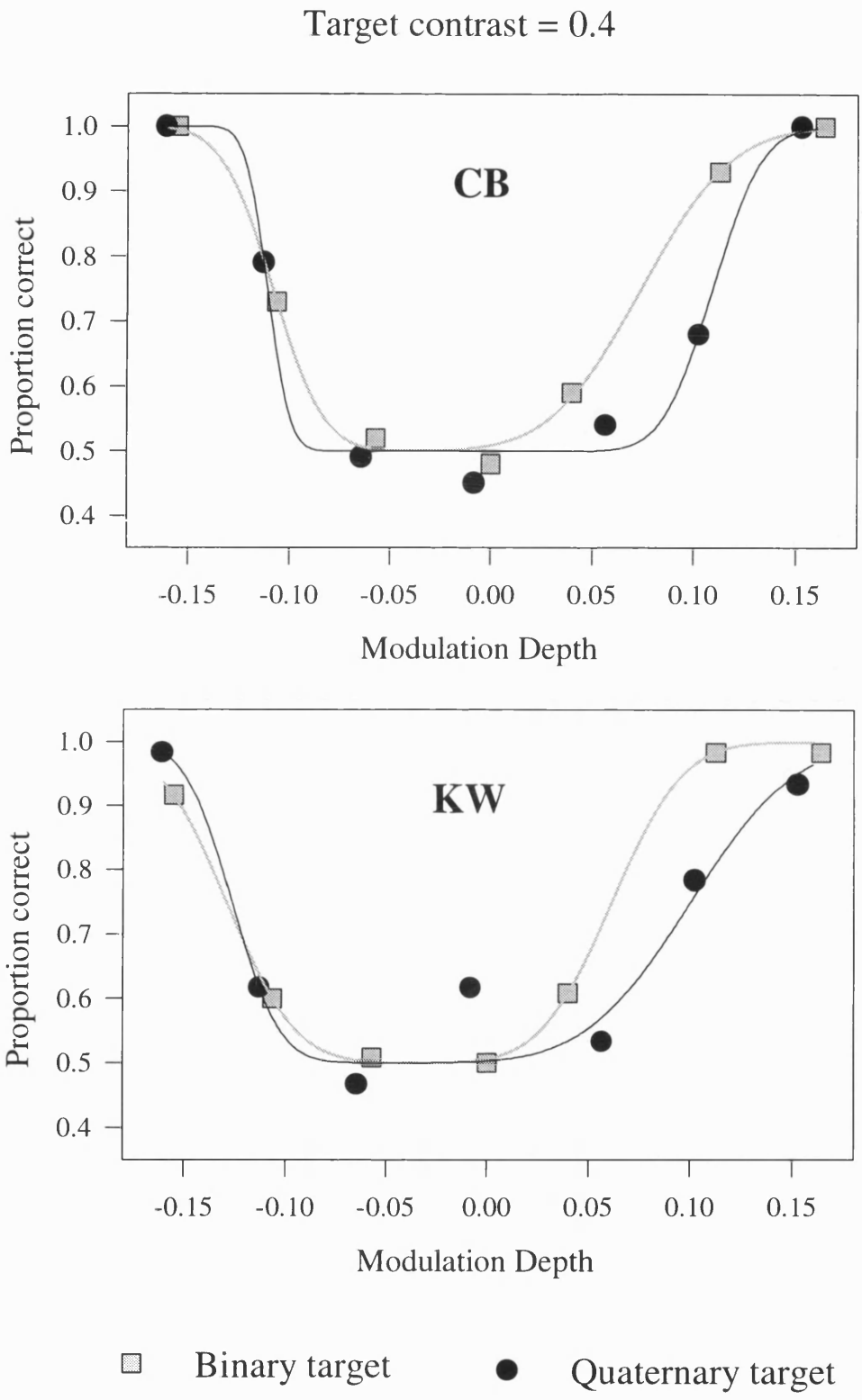


Figure 6.8: Results from experiment 6.2 showing modulation depth tuning curves for the detection of motion direction. The grey line shows the result of fitting equation 6.4 to the data derived from stimuli with standard binary targets. The black line shows the result of fitting the curve to the data derived from stimuli with quaternary targets.

tracking process of the kind outlined by Cavanagh (1991). If such a process was able to track luminance microstructure defined motion, one would not expect the performance on stimuli containing microstructure defined motion to drop to chance when the stimuli contained no contrast defined motion.

The lack of an effect of luminance microstructure supports the predictions drawn from the second-order channel hypothesis. Initially these results also appear to provide evidence against the idea that an early non-linearity is responsible for our perception of second-order motion. With the widely different luminance values present in the quaternary noise compared to the binary noise one might expect that a non-linearity applied before bandpass filtering would result in different levels of mean output after some non-linear transformation. If the perception of contrast defined motion were dependent upon an early non-linearity, one might therefore expect to see a shift in the curves between standard binary and quaternary targets. One can see that this is not necessarily the case.

It is assumed that some non-linear transformation $T(x)$ is applied to the signal before bandpass filtering. Effectively this can be seen as applying a non-linearity directly to the input luminance and might possibly reflect the non-linear relationship between photoreceptor input and output. The non-linear function may be expressed as a power function

$$T(x) = \sum_{n=1}^{n=\infty} k_n . x^n \qquad (6.5)$$

(Brown & Glazier, 1980). Early non-linearities in human vision have been analysed in these terms by taking the simple step of truncating the series so that the non-linearity is represented only by the first and second terms (Derrington, 1987). Generally, as the order of the terms increase, the contribution of each successive term is seen to decrease (Derrington, 1990).

In the stimuli investigated above, there was no systematic shift in the curves for standard binary and quaternary targets. For the early non-linearity hypothesis to be sustained, it would have to be supposed that the mean levels after non-linear transformation of both noise types, were very similar. In other words, if the RMS contrast and mean luminance of the two noise patterns are the same, then the mean level after some non-linear transformation

must also be approximately equal. In fact, if we use only the first two terms of the power series, we can show that the mean level after transformation is dependent only upon contrast and mean luminance.

The luminance value I_n of any random dot pattern containing h luminance levels each with a probability $P(I_n)$ of occurring can be written as $I_{mean} + x_n$. Because the pattern averages to mean luminance we know that

$$\sum_{n=1}^{n=h} (I_{mean} + x_n).P(I_n) = I_{mean}$$

$$\therefore \qquad (6.6)$$

$$\sum_{n=1}^{n=h} x_n.P(I_n) = 0 \ ,$$

and of course

$$\sum_{n=1}^{n=h} P(I_n) = 1 \ . \qquad (6.7)$$

From the equation for calculating RMS contrast (equation 6.3) we can also show that

$$\sum_{n=1}^{n=h} x_n^2.P(I_n) = (I_{mean}.C_{RMS})^2 \ . \qquad (6.8)$$

If we apply the non-linear transformation $T(x) = ax + bx^2$ to each luminance level and then take the average (also taking into account the probability of a particular luminance level occurring) then the mean output X can be written as follows:

$$X = a.\sum_{n=1}^{n=h} (I_{mean} + x_n).P(I_n) + b.\sum_{n=1}^{n=h} (I_{mean} + x_n)^2.P(I_n) \ , \qquad (6.9)$$

which can be simplified to

$$X = a.I_{mean} + b.\sum_{n=1}^{n=h} (I_{mean} + x_n)^2.P(I_n) \ , \qquad (6.10)$$

then expanded

$$X = a.I_{mean} + b.\sum_{n-1}^{n-h} I_{mean}^{2}.\mathrm{P}(I_n) + 2b.\sum_{n-1}^{n-h} I_{mean}.x_n.\mathrm{P}(I_n) + b.\sum_{n-1}^{n-h} x_n^{2}.\mathrm{P}(I_n) \ , \qquad (6.11)$$

from which we can see that

$$X = a.I_{mean} + b.I_{mean}^{2} + b.(I_{mean}.C_{RMS})^2 \ . \qquad (6.12)$$

This shows that, using the function $T(x) = ax + bx^2$ as the proposed non-linear function, then whatever the values of a and b, the mean level after transformation of a patch of noise is independent of the luminance microstructure. It is dependent only upon the mean luminance and the root mean square contrast of the pattern. Therefore, using the linear and quadratic terms of the power series expansion, the distortion product hypothesis also predicts that noise patterns of the same root mean square contrast but different luminance microstructures should be interchangeable.

However, as previously noted, the use of only the first two terms in the power series expansion is an approximation. The more terms we include, then the more accurate the approximation becomes. If we do wish to include additional terms then it is clear that two patches of noise with the same RMS contrast will not necessarily have the same mean level after the non-linear transformation. In order to test the distortion product hypothesis we need to devise patterns that can have identical RMS contrasts but very different luminance distributions. Asymmetric noise distributions may provide such an opportunity. In the next five experiments, the effects of using asymmetric or "biased" noise are investigated (the luminance compositions of the patterns are described in the general methods section).

In these experiments "top" and "bottom" biased noise are employed. The terms "top" and "bottom" are used to describe which region of the noise has the greatest luminance range when the luminance values are considered as perturbations from mean luminance. For each particular noise type, the top and bottom biased patterns can have very different luminance ranges but the same contrast. Furthermore, since the top and bottom patterns within each noise type are essentially reflections of one another through mean luminance, the expected Fourier spectra of the two patterns have to be exactly the same. We can think of the pattern in two ways, either as a set of deviation from mean luminance or as the sum of its Fourier

components added to mean luminance. If we multiply the deviations from mean luminance by minus one then, to obtain the reflected pattern, we can also multiply the Fourier components by minus one. This is equivalent to simply shifting the phase of each component by half of its cycle.

Whilst it seems reasonable to suggest that the output of a system based upon linear filtering followed by fullwave rectification or squaring should be unaffected by the pattern polarity, it would seem intuitive that halfwave rectification applied to the output of bandpass linear filters should be dependent upon the stimulus polarity. Indeed, the use of stimulus polarity might seem a powerful paradigm with which to examine whether halfwave mechanisms are implicated in the perception of second-order motion.

In Figure 6.9, the outputs of a two channel model with different post-filtering non-linearities are shown. A Laplacian of a Gaussian spatial filter was applied to space-time plots of static pattern motion sequences with biased binary noise targets. As in the experiments detailed in this chapter, the comparison pattern was standard binary noise. The width of the noise patches was 32 times that of the noise. The noise had a width of 4 pixels. The profile of the spatial filter relative to the size of the noise checks is shown in Figure 6.10. Each input image was 640 by 640 pixels. Results are drawn from the central 512 by 512 region of the processed input image. The Fourier amplitude spectra of 400 instantiations of each level of stimulus were added together. The output was calculated as the difference between the summed amplitude at the absolute modulation spatial and temporal frequencies moving in the direction of the modulant and the summed amplitude at the absolute modulation spatial and temporal frequencies moving in the opposite direction to the modulant. It can be seen that, whatever the stimulus polarity and whichever of the post filtering non-linearities are used, there are no systematic differences between the curves for stimuli containing top biased binary targets and bottom biased binary targets.

However, when non-linearities are applied directly to the luminance surface, then we can see that there are substantial differences between the response to the differently biased patterns. In Figure 6.11, a compressive non-linearity of the form $T_1(I_n) = \log_e(I_n)$ and an expansive non-linearity of the form $T_2(I_n) = (I_n)^{1.5}$ were applied to the luminance values present in the images described in the paragraph above. Mean luminance was arbitrarily set to 14.36 (no

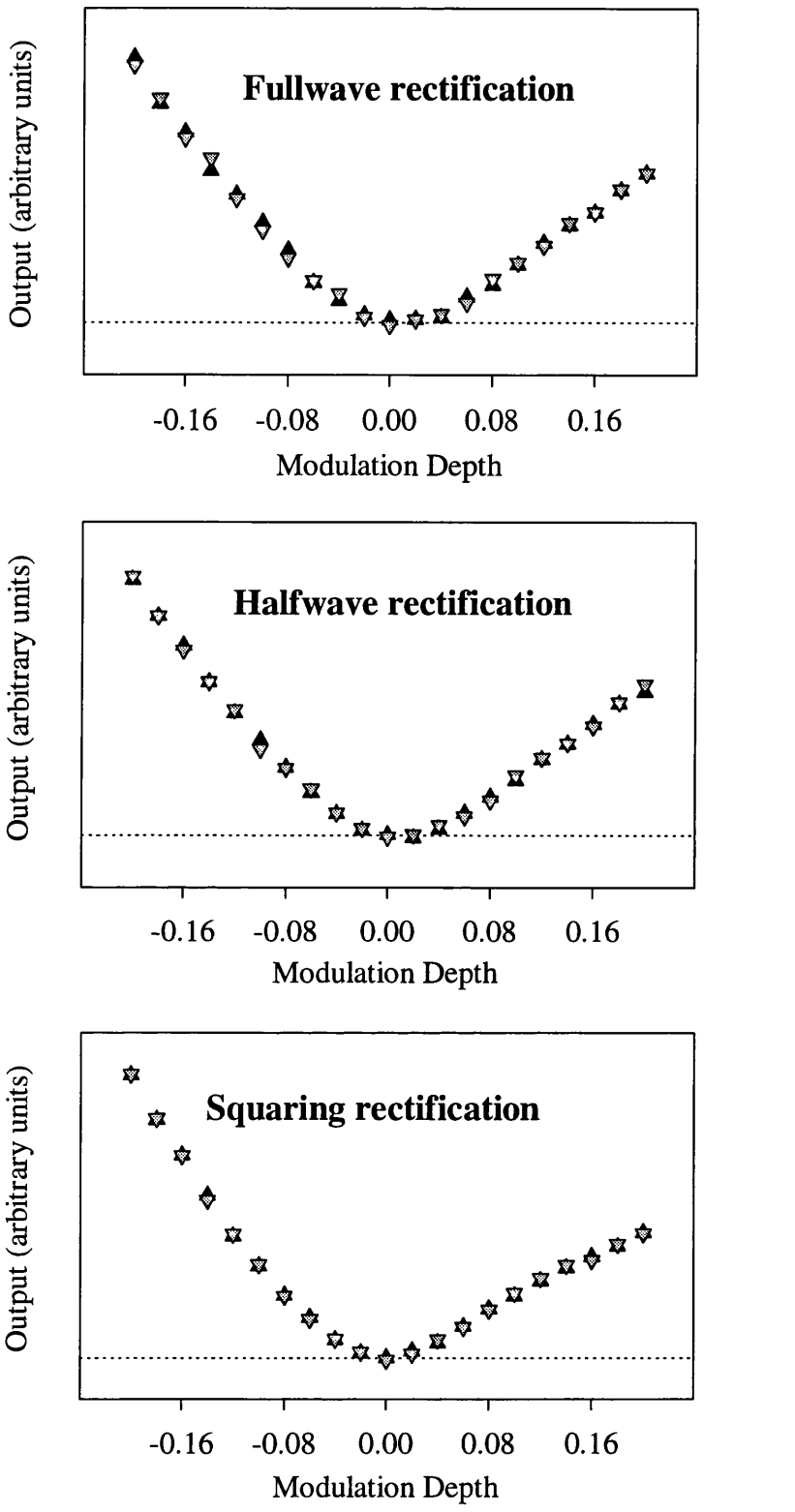


Figure 6.9: Results of applying fullwave, halfwave and squaring rectification to bandpass filtered images expressed as a function of modulation depth. See text for details.

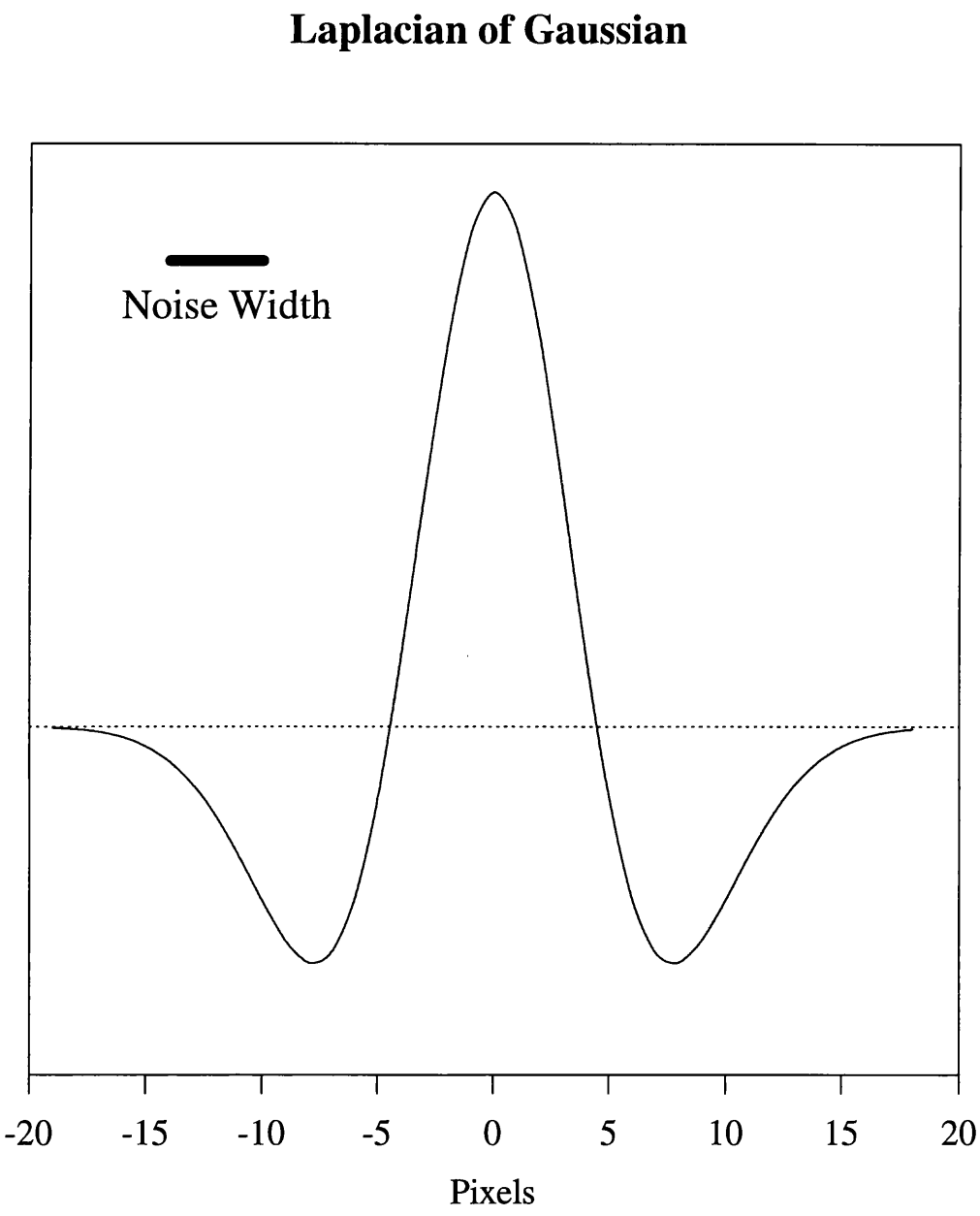


Figure 6.10: Laplacian of Gaussian spatial filter. This was applied to images containing contrast and/or luminance microstructure defined motion. The width of the noise checks relative to the filter is shown by the bar in the top left hand corner of the graph.

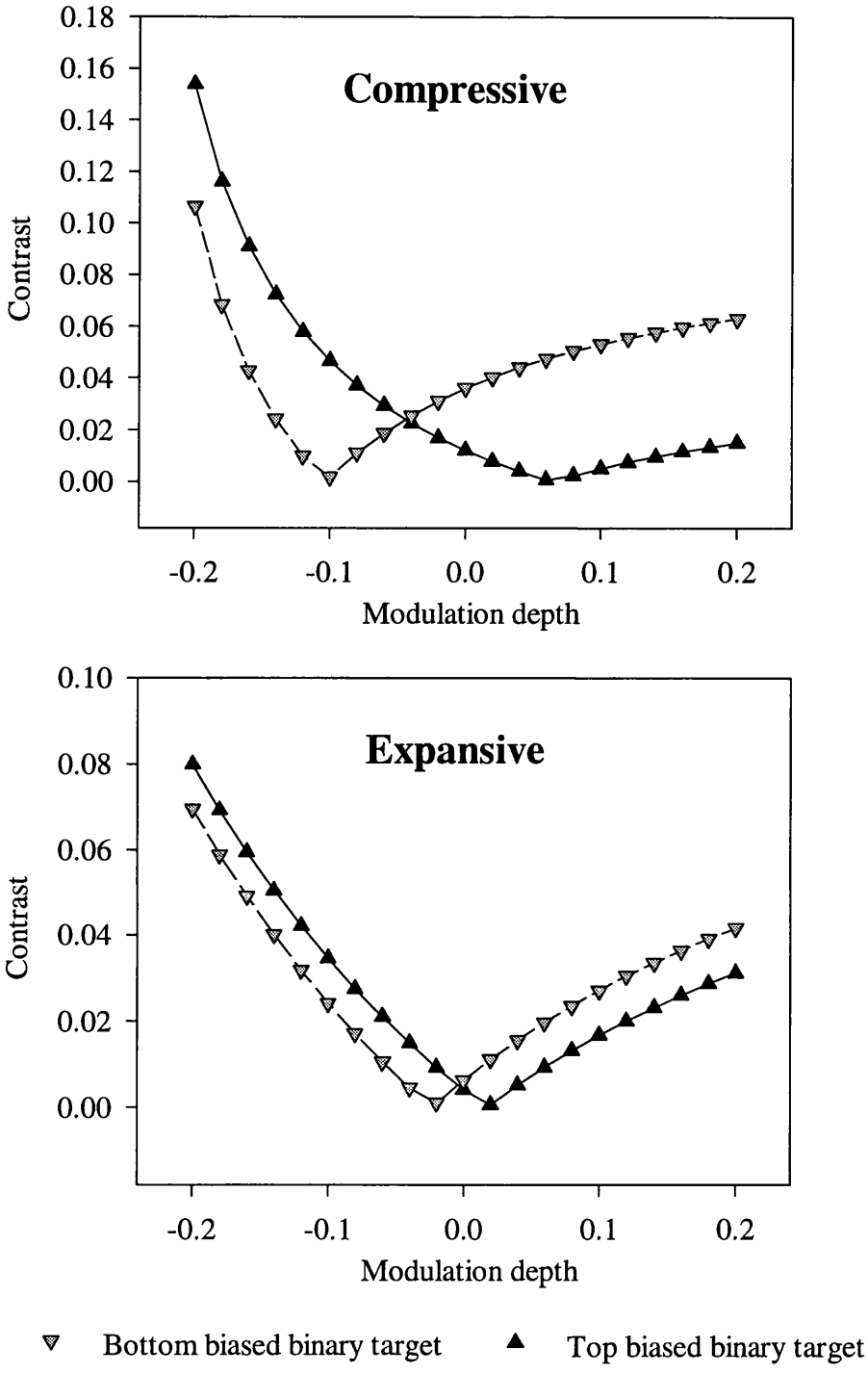


Figure 6.11: Contrast between expected mean levels of constituent noise patterns in images containing motion defined by contrast and/or luminance microstructure.

units) with a notional maximum luminance of 28.72 and a minimum luminance of zero. This ensures that, as in the real world, all luminance values are positive. Rather than generate a series of images and take Fourier spectra, a measure of signal strength introduced by the non-linearities can be calculated simply from the mean levels after non-linear transformation within the comparison and target patterns. The mean level within each pattern may be calculated by the following equation:

$$T_{mean} = \sum_{n-1}^{n-h} P(I_n).\mathrm{T}(I_n) \qquad (6.13)$$

where the function T(x) is the non-linear transformation. The contrasts shown in Figure 6.11 are RMS contrasts calculated from the mean post-transformation levels for the target noise and comparison noise. As the modulant is a square wave, the RMS contrast between the mean levels is identical to the Michelson contrast. With both of the non-linearities examined in these simulations, there is a clear shift between the curves for the bottom pattern and the top pattern. The curve for the top biased noise target is shifted to the right, whilst the curve for the bottom biased target is shifted to the left. This means that the bottom biased target effectively has the higher "contrast". In fact the non-linearities have been chosen to produce these shifts. As we shall see, the relative shift between the two curves is dependent upon the particular non-linearity.

If we apply the non-linear transformation $\mathrm{T}(x) = ax + bx^2 + cx^3$ to each luminance level of a noise pattern, the mean output X may be written as follows:

$$X = a.\sum_{n-1}^{n-h} (I_{mean} + x_n).\mathrm{P}(I_n) + b.\sum_{n-1}^{n-h} (I_{mean} + x_n)^2.\mathrm{P}(I_n) + c.\sum_{n-1}^{n-h} (I_{mean} + x_n)^3.\mathrm{P}(I_n) \quad . \qquad (6.14)$$

From equation 6.12 we can rewrite the first two terms to give

$$X = a.I_{mean} + b.I_{mean}^{\ 2} + b.(I_{mean}.C_{RMS})^2 + c.\sum_{n-1}^{n-h} (I_{mean} + x_n)^3.\mathrm{P}(I_n) \quad . \qquad (6.15)$$

If we take the cubic term by itself and expand, this gives

$$c.\sum_{n-1}^{n-h}(I_{mean}+x_n)^3.P(I_n) = c.\sum_{n-1}^{n-h} I_{mean}{}^3.P(I_n) + 3c.\sum_{n-1}^{n-h} I_{mean}{}^2 x_n.P(I_n) + 3c.\sum_{n-1}^{n-h} I_{mean} x_n{}^2.P(I_n) + c.\sum_{n-1}^{n-h} x_n{}^3.P(I_n) , \quad (6.16)$$

which can then be simplified to

$$c.\sum_{n-1}^{n-h}(I_{mean}+x_n)^3.P(I_n) = cI_{mean}{}^3 + 3cI_{mean}{}^3.C_{RMS}{}^2 + c.\sum_{n-1}^{n-h} x_n{}^3.P(I_n) . \quad (6.17)$$

Substituting equation 6.17 back into equation 6.15 gives

$$X = aI_{mean} + bI_{mean}{}^2.(1 + C_{RMS}{}^2) + cI_{mean}{}^3.(1 + 3.C_{RMS}{}^2) + c.\sum_{n-1}^{n-h} x_n{}^3.P(I_n) . \quad (6.18)$$

From equation 6.18 we can see that the inclusion of the cubic term in the power series expansion means that X includes a term that is dependent upon luminance microstructure. To take the example of top biased noise and bottom biased noise. If both patterns have the same mean level and the same contrast, then the values of X for each pattern will differ only by the cubic term in equation 6.18. The pattern that acts as if it has the higher contrast will be the one that gives a greater value of X. If the top biased binary noise has the format ([+2, ⅓], [-1, ⅔]) then, putting these values into the cubic term in equation 6.18 gives a total of $c.4⅔$. In the case of the bottom biased pattern, this total is $-c.4⅔$. If c is positive, then the top biased pattern will effectively have the higher contrast but if c is negative then the bottom biased pattern will effectively have the higher contrast.

Therefore, by comparing modulation depth tuning curves for direction discrimination in stimuli with top biased targets to those with bottom biased targets, we may have a method of assessing whether non-linearities prior to bandpass filtering may influence the perception of motion direction in contrast defined motion sequences. If there is a consistent bias in the sign of c then the two curves should be shifted with respect to one another.

6.3.2.1. Experiment 6.3: Asymmetric tertiary noise - dynamic patterns.

This experiment contained only one set of trials. The target in one of the two runs was top biased tertiary noise, the target in the other run was bottom biased tertiary noise. Target contrast was 0.6. For both runs the comparison pattern consisted of standard binary noise. All stimuli contained dynamic patterns. The data for the two subjects are plotted in Figure 6.12. Subject CB shows a clear difference between the two target patterns with the U-shaped curve for the top biased pattern shifted to the right and that for the bottom biased pattern shifted to the left. It should be noted that, on the left hand side of both curves, very few data points contribute towards the regression. Most of the data points are either close to the 100% and 50% correct marks and therefore contribute little information about the precise form of the curve. The results from subject CB would appear to be supported to some small extent by those from subject KW who shows shifts in the same direction but with reduced amplitude and wider U-shapes. Because of the reduction in amplitude, these results must be treated with some degree of caution. The gamma correction function for the experimental set up was derived through linear regression in log-log space. The fit will be better constrained in the centre of the range than at the edges. With a target contrast of 0.6, the contrast of the comparison noise has to be quite high for reasonably large negative modulation depths to be obtained. Therefore with negative modulation depths we might possibly expect to find some inaccuracies. However, subject CB shows the largest shift on the left hand sides of the two curves which would argue against this analysis.

6.3.2.2. Experiment 6.4: Asymmetric tertiary noise - static patterns.

The experiment contained one set of trials. As in experiment 6.3, top biased and bottom biased tertiary noise targets were used. In this experiment, the target contrast was 0.3 and the stimuli contained static patterns. The data from the two subjects are plotted in Figure 6.13. Neither of the subjects show any substantial differences between the two conditions.

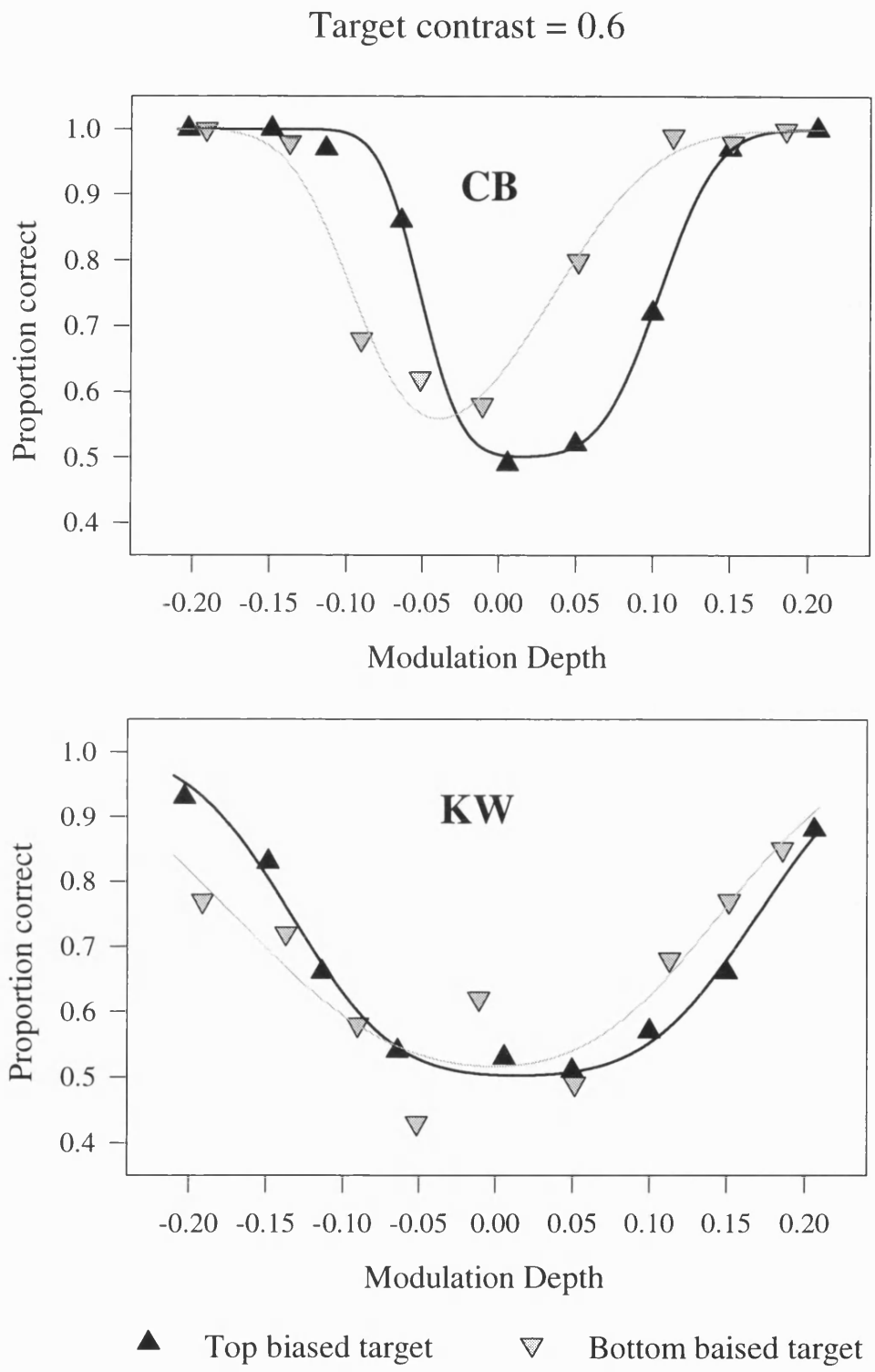


Figure 6.12: Results from experiment 6.3 showing modulation depth tuning curves for the detection of motion direction. The grey line shows the result of fitting equation 6.4 to the data derived from stimuli with bottom biased tertiary targets. The black line shows the result of fitting the curve to the data derived from stimuli with top biased tertiary targets.

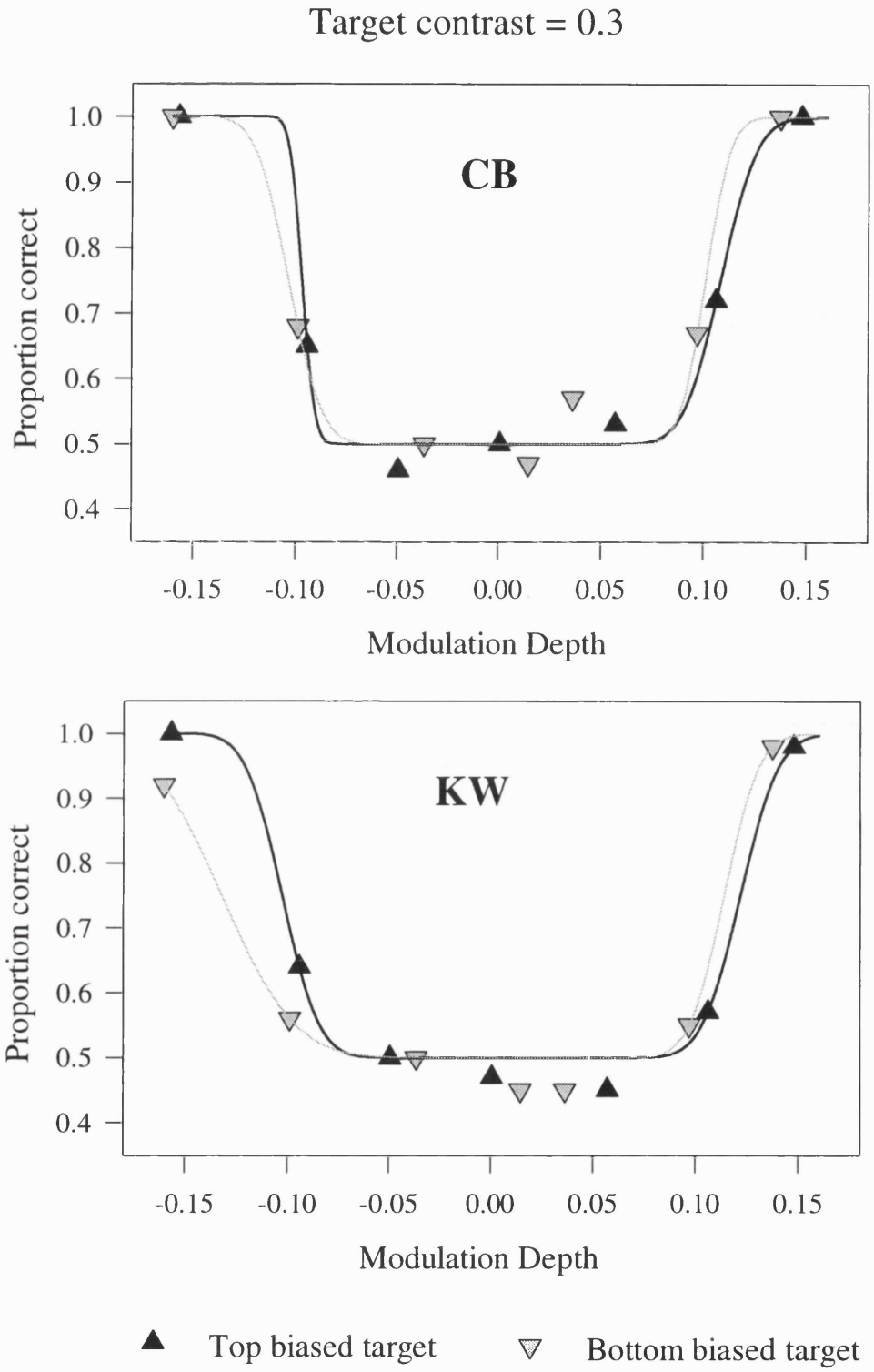


Figure 6.13: Results from experiment 6.4 showing modulation depth tuning curves for the detection of motion direction. The grey line shows the result of fitting equation 6.4 to the data derived from stimuli with bottom biased tertiary targets. The black line shows the result of fitting the curve to the data derived from stimuli with top biased tertiary targets.

6.3.2.3. Discussion of results from Experiments 6.3 & 6.4.

If second-order motion were perceived through the operation of an early non-linearity, one might well expect some shift to occur between modulation depth tuning curves for direction discrimination with top biased and bottom biased target noise patterns. If no such difference is found then this provides supporting evidence for the notion of a second-order channel. In Experiment 6.3, subject CB appears to show a large shift between the two conditions. Subject KW shows a similar pattern of response but with far less difference between the two conditions. In Experiment 6.4, there is little evidence for any difference between the two conditions.

It might seem reasonable to conclude that the pattern of results obtained from subject CB in experiment 6.3 are in some way anomalous. However, it is difficult to see how this anomaly may have arisen. One obvious line of reasoning is to question the gamma correction of the system; however the gamma correction was repeatedly tested and no significant deviations from linearity were found. Furthermore the shift between the two curves shown by subject CB is quite substantial. It would require that the calculated contrasts of the top and bottom biased patterns differ from one another by about 8%. It seems unlikely that such a difference could occur without being accompanied by some measurable luminance non-linearity.

If the results from subject CB in experiment 6.3 cannot be put down to some problem with the experimental setup, then it is clear that some genuine effect is being measured. This leaves the problem of how to account for the lack of such an effect in experiment 6.4. One possible reason for the difference between the two experiments lies in the difference in target contrast. If the results of experiment 6.3 reflect the operation of some early non-linearity, then we might assume that the contribution of the non-linearity increases as contrast increases. If this were the case then the effect of the non-linearity would be smaller in experiment 6.4 than in experiment 6.3.

If there is an effect, then it is clear that it is small, at least in subject KW. The experimental set up in experiments 6.1 through to 6.4 suffers from the problem that the graphics card only has 8 bit output. As described in the method section, a procedure to reduce the possibility

of luminance contamination had to be incorporated during the creation of the stimuli. It was because of this that predetermined levels were used rather than some adaptive procedure. The results from many of the experiments reflect this problem in that the modulation depth tuning curves are often sampled inefficiently with a large number of trials placed at regions where the response of the subjects is either close to 100% or is close to chance. It is obvious that the use of some adaptive procedure would be much better suited to this task. However, if the step sizes are to be small and accurate enough so that an adaptive procedure can be used, then the experiment must be repeated on a system with a larger number of output levels.

6.3.3.1. Experiment 6.5: Asymmetric binary noise - static patterns.

(a). Method.

Stimuli were displayed on a Manitron monitor driven by a Cambridge Research Systems VSG 2/3 graphics board. This board had a DSP "grating generator" capable of producing 14 bit output (the information on the DSP is also of 14 bit resolution). Gamma correction was carried out using a Graseby (UDT) model 265 luminance probe. Linearisation was achieved through the use of a look-up table. The grating generator is designed primarily to produce gratings, not to produce spatially two dimensional random noise stimuli. Therefore a facility to window the output of the grating generator using the graphics card was utilised. The grating generator can supply 4 outputs (3 gratings and a background). All of the outputs can be windowed using the graphics card. If a particular pixel on the screen is represented by an area of graphics memory, that area of graphics memory may be used to window the output from the grating generator. Any particular pixel on the screen can take one of the four values of the outputs from the grating generator. Which output is displayed at that location depends upon the value in graphics memory that corresponds to the location on the screen.

By fixing the DSP outputs to a constant value and by cycling through images in graphics memory, random dot images with 14 bit output can be produced. The use of this experimental setup meant that it was only possible to display 4 luminance values on the screen at any one time. Therefore the stimuli were displayed over the entire screen. This in

turn meant that the images used were large enough for considerable time to be taken to generate a fresh image. The images were therefore generated and stored off-line. During the experiment they were loaded into graphics memory from a "RAM disk", a disk created in computer memory. Because the images in graphics memory window the grating generator output, the 4 values in the image can represent any luminance value that he system is capable of producing. The arrangement used here can be conceptualised more simply if the grating generator is thought of as a look-up table which can contain only 4 values.

As static patterns were utilised, only 4 whole screen images (stimulus frames) were needed for each motion sequence. However, because the images needed to be generated off-line, and because the storage on the RAM disk was limited, only 12 different motion sequences were utilised in this experiment (ie. basically all the same sequence but with different instantiations of underlying noise). For each trial in the experiment, one sequence was randomly chosen. Direction of motion, leftwards or rightwards, was randomised. The direction could readily be reversed simply by stepping through the stimulus frames in reverse sequence. The starting position was randomised by randomly choosing the start frame. For each stimulus, the 4 stimulus frames were cycled through twice, each frame remaining on the screen for 4 physical frames to give a temporal frequency of 3.75 Hz.

Other than the methods for stimulus generation and display, the stimuli themselves were very similar to those from experiments 6.2 and 6.4. However, because only 4 luminance values were available, two for the target and two for the comparison, biased binary noise was used as the target noise pattern. The viewing distance was 1.5 metres; the display area of the screen and the stimuli themselves had a width of 13.76° and a height of 9.79°. The noise checks were 2.53 arc minutes wide and 2.59 arc minutes high. Modulant spatial frequency was 0.22 cycles per degree meaning that each noise patch was 2.28° wide (and extended for the full height of the screen). The mean luminance was 19.56 cd/m^2. There was a minimum gap of 2 seconds between trials.

A slightly modified form of the staircase procedure PEST (Parameter Estimation by Sequential Testing) described by Taylor & Creelman (1967) was utilised to measure the modulation depth tuning curves. As suggested by Findlay (1978), the cumulative results of the trial were fitted to a psychometric function using probit analysis (Finney, 1971). Each

run contained 64 trials. The run finished when all the trials had been responded to rather than when the step size reached a certain value. Subjects were presented with a set of trials containing three interleaved runs. All of the runs within a set had the same target noise type. Both of the target types had contrasts of 0.6. Within each set, one of the runs was set to hone in on the 75% mark, the other two were set to hone in on the 84% and 66% marks (ie. one standard deviation on either side of the 75% mark). In this case, PEST is simply being used to provide a scattering of stimulus levels that are strung out along the length of the psychometric function. This procedure should mean that the numbers of trials presented at levels where subjects performance is either 100% or 50% should be minimised.

One problem with the procedure described so far is that we are attempting to measure U-shaped curves. This could potentially confuse the operation of a staircase procedure, particularly if the step size become large relative to the width of the U-shape. To evade this problem, the maximum step size was kept small. Subjects were given 4 sets of trials, one for each side of the function for each target. These 4 sets of trials were repeated 4 times for subject CB and 2 times for subject KW. Before running the sets from which the data was collected both subjects had run through at least two sets of the 4 sets of trials. Subjects indicated the perceived direction of motion by pressing the left/right arrows of a PC keyboard. Feedback was provided to let subjects know whether their previous response had been correct or incorrect. A fixation point was provided in the centre of the screen. Subjects were required to fixate during each stimulus presentation.

(b) Results.

The results for the two subjects are shown in Figure 6.14. There are fewer data points for subject KW because the minimum step size was much larger than that used by subject CB and because twice as much data was collected for subject CB. There are therefore more trials per data point for that subject. All points with less than 10 trials were excluded from the probit analyses and from the graphs. The data obtained in this experiment show exactly the same pattern as that shown in experiment 6.3. In the present experiment the magnitude of the difference between the top biased and bottom biased patterns is far larger for both subjects. The U-shaped curve for the bottom biased pattern is shifted to the left whilst that for the top biased pattern is shifted to the right. It would seem, within the confines of this

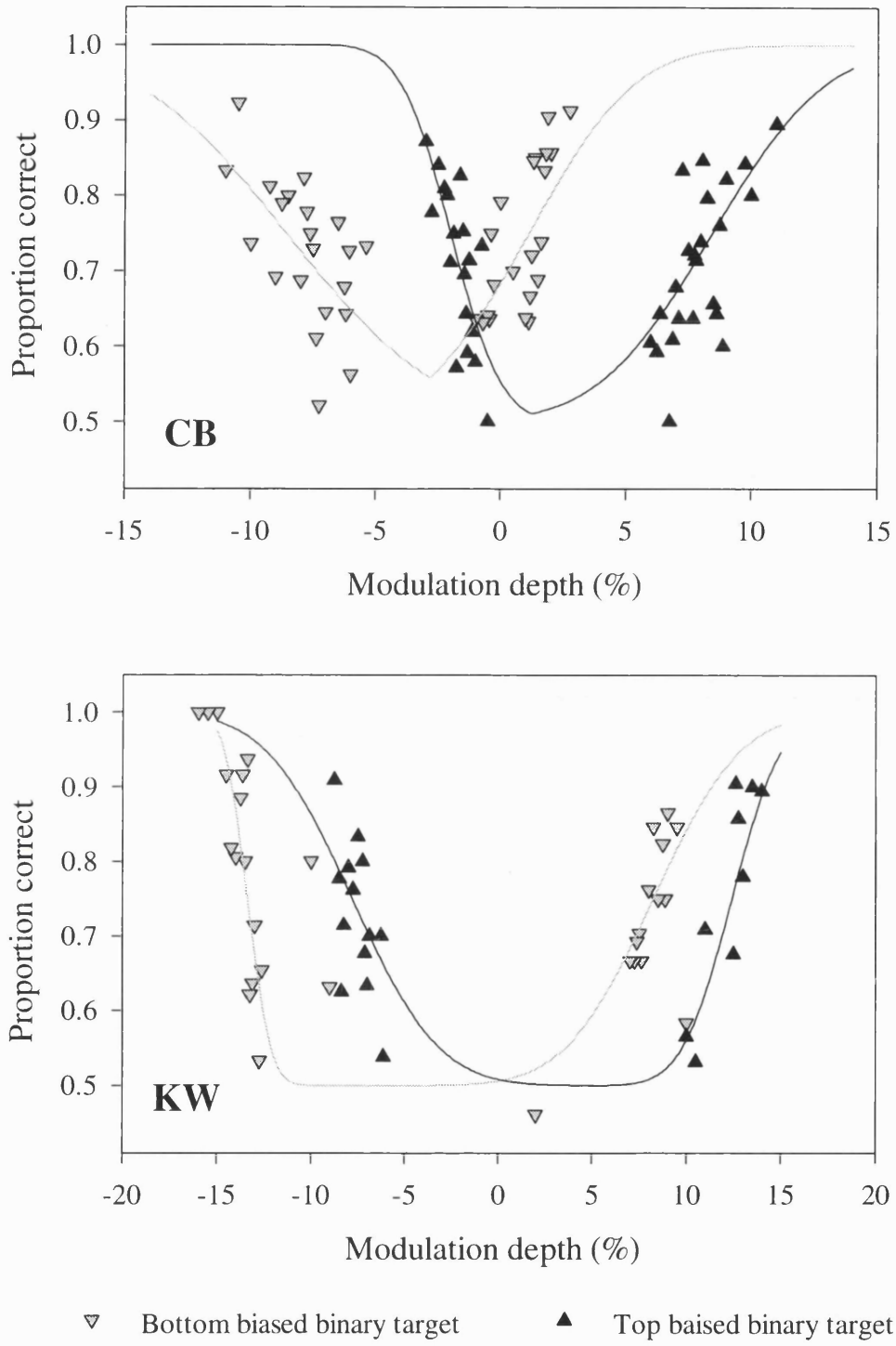


Figure 6.14: Results from experiment 6.5 showing modulation depth tuning curves for the detection of motion direction. The grey line shows the result of fitting psychometric functions to the data derived from stimuli with bottom biased tertiary targets. The black line shows the result of fitting psychometric functions to the data derived from stimuli with top biased tertiary targets.

particular task, that the bottom biased pattern effectively has a higher contrast than standard binary noise of the same RMS contrast whilst the top biased pattern effectively has a lower contrast.

6.3.3.2. Discussion of results from Experiment 6.5.

The results from experiment 6.5 show a difference between contrast defined motion sequences containing bottom biased and top biased binary noise. It appears that the bottom biased binary noise target effectively has a higher contrast than the top biased binary noise target. This finding cannot be explained on the basis of bandpass linear filtering followed by some rectifying non-linearity, but may be accounted for by appealing to some early pre-bandpass filtering non-linearity. If the non-linearity is described by the equation $T(x) = ax + bx^2 + cx^3$, then the value of c must be negative (see equation 6.18).

Chubb *et al* (1994), examining contrast defined motion with symmetrical manipulations of luminance microstructure, also obtained results that implied the operation of some process other than bandpass filtering followed by rectification. They proposed a non-linearity of the form, $r(c) = \max\{0, f(|c| - \tau)\}$ where c is local contrast, τ is some threshold and the function $f(x)$ is monotonic and increases as x increases. For values of c that are less than τ, a value of zero is produced. It is unclear whether Chubb *et al* believe that this non-linearity exists as an early non-linearity in the first-order channel or whether they believe that this non-linearity constitutes a separate second-order channel. However the contrast transforming non-linearity might well be modelled by an early luminance non-linearity and the data produced by Chubb *et al* may therefore also be seen to provide supporting evidence for the distortion product hypothesis. The source of such a non-linearity may well lie in the process of phototransduction. If this is true, then a similar effect might well be present in spatial tasks. The next two experiments examine whether or not this is the case.

6.3.4.1. Experiment 6.6: Contrast comparison - asymmetric binary noise.

Subjects compared the contrast of two patches of static noise and indicated which of the two had the higher contrast. One of the patches consisted of top biased binary noise, the other of bottom biased binary noise. The stimuli were generated and displayed on a Sun

Sparcstation LX, the display of which had been carefully gamma corrected using a Graseby (UDT) model 265 luminance probe. Each of the 8 subjects was presented with 11 trials. In each trial the average contrast across both noise patches was 0.5. However, both patches could differ from mean contrast by the same amount but in opposite directions. So if the contrast of the top biased patch was 0.6 then the contrast of the bottom biased patch was 0.4. The stimulus can therefore be characterised in terms of the contrast difference between the two patches where the contrast difference is equal to the contrast of the top biased pattern (C_{top}) minus the contrast of the bottom biased pattern (C_{bottom}). The 11 trials presented to each of the 8 subjects had contrast differences running from -0.20 to 0.20 in steps of 0.04.

The monitor was only capable of 8 bit output. To guarantee that mean luminance was identical in both patches, discreet output levels were assigned in the following manner. To take a concrete example, suppose a luminance of 20.4 cd/m^2 was required. The nearest grey levels on either side would be calculated. A grey level of 135 gives a luminance value of 20.2678 cd/m^2 whilst a grey level of 136 gives a luminance value of 20.5653 cd/m^2. A random procedure is used to choose between the two possible grey levels. This is done in such a way that, with an infinite number of random choices, the mean calculated luminance from all the grey levels chosen by the procedure would be equal to the required luminance (see Chapter 2). This procedure was also employed in Experiment 6.7.

The two patches were presented horizontally adjacent to one another. Subjects were asked to seat themselves at a comfortable "working distance" from the screen (ie. about 75 cm ±10 cm). Each patch was 8.82° on each side (from a distance of 75cm). The screen on which they were presented was 26.05° wide by 20.41° high. The noise check size was 4.14 arc minutes. Mean luminance was 37.4 cd/m^2. Outside the display area, the screen was set to mean luminance. The side on which each pattern type was presented and the order of the trials were randomised. Subjects indicated which patch they thought had the highest contrast by saying "right" or "left". The responses were recorded by the experimenter. All eight subjects were workers in the field of vision research who had frequently participated in psychophysical experiments. All had normal or corrected to normal vision.

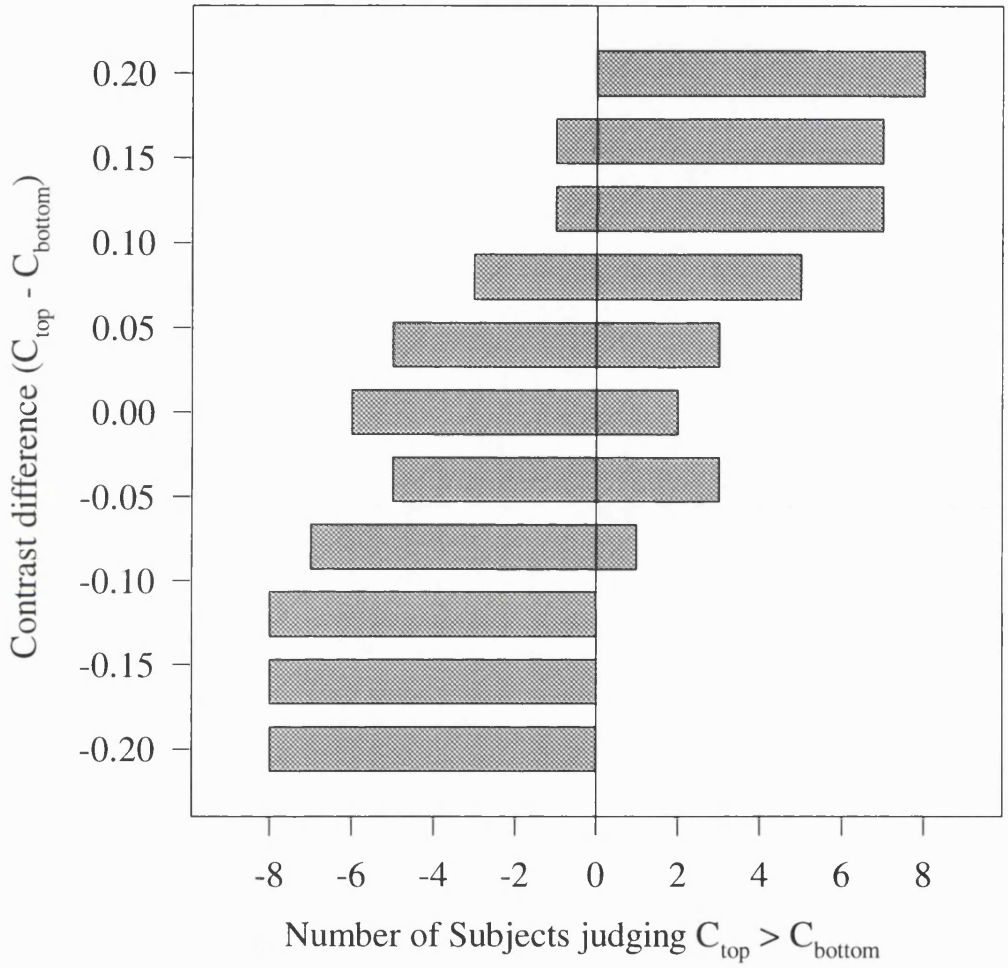


Figure 6.15: Results from experiment 6.6 showing number of subjects (out of 8) judging the contrast of top biased binary noise as greater than that of bottom biased binary noise. Negative numbers show subjects indicating that the contrast of the bottom biased pattern appeared to be greater than that of the top biased pattern.

Results collapsed across subjects are shown in Figure 6.15. It is clear that there is a tendency to judge the bottom biased noise as having a higher contrast that the top biased noise, even when the latter may have a higher RMS contrast. These results support those obtained in experiment 6.5. If perceived contrast is indeed a measure of the amplitude output of bandpass filters, then clearly this is yet more evidence for some non-linearity preceding bandpass filtering.

6.3.4.2. Experiment 6.7: Contrast Matching - asymmetric binary noise.

Experiment 6.6 provided evidence that bottom biased noise has a higher perceived contrast than top biased noise. The evidence would be more convincing if the effect could be quantified. In this experiment the contrast of a standard binary noise comparison patch is matched to that of target patches of top and bottom biased noise. Additionally in experiments 6.3 and 6.5 there was evidence of a shift between the modulation depth tuning curves for top biased and bottom biased noise. In experiment 6.4 this was not the case. One simple explanation for the lack of such an effect might be that the size of the proposed distortion product scales with contrast. If this is the case, and if perceived contrast reflects the operation of a distortion product, then we ought to find that the proportional difference between perceived contrast and RMS contrast ought to decrease as contrast decreases.

Stimuli were generated and displayed on a Sun SparcStation LX. Gamma correction and mean luminance were identical to those described in experiment 6.6. The viewing distance was 75cm. The task of the subjects was to indicate which of two patches of noise had the higher contrast by depressing a mouse button. The stimulus consisted of a rectangle with a square of target noise on one side and a square of the comparison noise on the other. Each square was 5.89° on each side, the squares were separated from one another by a region of mean luminance with a contrast of zero and a width of 0.23°. The noise check size was 2.76 arc minutes. Each stimulus was presented for 2 sec, the minimum interstimulus interval was 1 sec. During the experiment the screen around the display rectangle was set to minimum luminance (0.4 cd/m^2). In the interstimulus interval the display rectangle reverted to mean luminance.

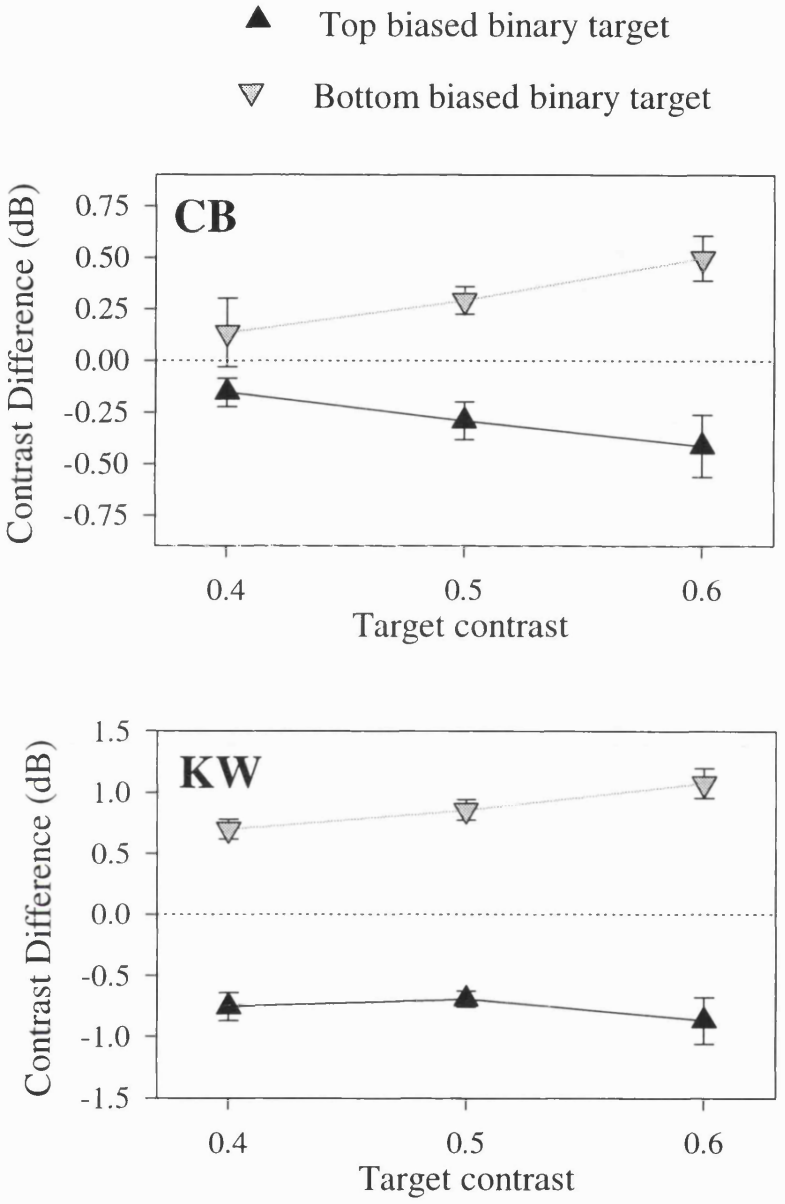


Figure 6.16: Results from experiment 6.7 (a contrast matching task) showing contrast difference as a function of target contrast. Contrast difference is calculated as 20 times the log of the ratio of the matching contrast and the target contrast.

The contrast of the target noise was held constant during each run. Each run contained 64 trials. The side on which the target pattern was presented was randomised for each trial. Within each run adaptive probit estimation (see Chapter 2) was used to find the point of subjective equality (PSE) where standard binary noise appeared to have the same contrast as the target. The target type was either top biased binary noise or bottom biased binary noise. Subjects were tested over three contrast levels, 0.6, 0.5 and 0.4. For each level and each type, four PSEs were extracted, giving a total of 24 runs. The order of presentation of the runs was randomised. Figure 6.16 shows the results for each of the two subjects in terms of the difference in log contrast between the mean PSE and the RMS contrasts of the target patterns. The results from experiment 6.6 are confirmed. Bottom biased noise appears to have a higher contrast and top biased noise appears to have a lower contrast than the RMS contrasts of the patterns. For subject CB, the proportional magnitude of this difference decreases quite sharply as contrast decreases. This trend is weaker in the results from subject KW.

6.3.4.3. Discussion of results from Experiments 6.6 & 6.7.

Studies examining perceived contrast using contrast matching tasks have generally been held to show that perceived contrast is linearly dependent on physical contrast (ie. Kulikowski, 1976; Ginsburg, Cannon & Nelson, 1980). The subjects in these studies have matched the contrasts of patterns that contain different spatial frequencies. Therefore the comparison between the two stimuli may not depend only upon the output of linear filters, but may also depend upon the way that these outputs are combined together (Quick, Hamerly & Reichert, 1976). Therefore where some non-linearity is proposed (Hamerly, Quick & Reichert, 1977), one cannot be certain to which process this applies. In experiments 6.6 and 6.7 the targets have exactly the same expected Fourier spectra. The data from these two experiments shows that the perception of contrast, measured by a simple contrast matching task, is subject to some non-linearity. If it is accepted that the perception of contrast is based upon the output of spatial frequency selective channels then the present study indicates that some non-linearity exists prior to bandpass spatial filtering.

6.3.4.3. Experiment 6.8: Estimating the necessary non-linearity.

As shown earlier, the shift between the modulation depth tuning curves for the top biased and bottom biased pattern must be due to those components in the power series expansion of the non-linear function that have a power of 3 or greater (see equation 6.12). However if we follow the approach of Derrington (1987, 1990), we can assume that the U-shaped tuning curves (rather than the shift between the biased noise targets) largely reflect the operation of the quadratic component. In the experiments described in this chapter, when comparing symmetric noise targets, there was little evidence for any effect of luminance microstructure. An effect was only obtained when comparing biased noise targets at a relatively high target contrast. If the modulation depth tuning curves described in this chapter are based upon the operation of an early non-linearity, the lack of an effect of noise type in a number of experiments implies that any early non-linearity is largely described by the operation of the quadratic component within the power series expansion. The equation $f(x) = ax + bx^2$ therefore provides a good approximation to the putative early non-linearity. This observation is useful because it allows us to make an estimate of the non-linearity necessary to produce the tuning curves detailed in this chapter.

From equation 6.3, modulation depth is defined as follows:

$$M = \frac{C_{target} - C_{comparison}}{C_{target} + C_{comparison}} . \qquad (6.19)$$

If a quadratic non-linearity, $f(x) = ax + bx^2$, operates on the luminance levels present in the target and comparison patterns then, from equation 6.12, we can work out the mean levels after the non-linear transformation as follows:

$$X_{target} = a.I_{mean} + b.I_{mean}^{2} + b.(I_{mean}.C_{target})^2 \qquad (6.20)$$

and

$$X_{comparison} = a.I_{mean} + b.I_{mean}^{2} + b.(I_{mean}.C_{comparison})^2 . \qquad (6.21)$$

Therefore after non-linear transformation, a stimulus with a modulation depth M containing noise of contrasts C_{target} and $C_{comparison}$ will contain a square wave first-order modulation with peaks and troughs of X_{target} and $X_{comparison}$ (or vice versa). In terms of the difference between these levels after transformation, the square wave will have a contrast of:

$$C_T = \frac{|X_{target} - X_{comparison}|}{X_{target} + X_{comparison}} . \tag{6.22}$$

By substituting equations 6.20 and 6.21 into 6.22 we obtain

$$C_T = \frac{|b.(I_{mean}.C_{target})^2 - b.(I_{mean}.C_{comparison})^2|}{2a.I_{mean} + 2b.I_{mean}^2 + b.(I_{mean}.C_{target})^2 + b.(I_{mean}.C_{comparison})^2} , \tag{6.23}$$

and dividing the top and bottom by $b.I_{mean}^2$ gives

$$C_T = \frac{|C_{target}^2 - C_{comparison}^2|}{k + C_{target}^2 + C_{comparison}^2} \quad \text{where} \quad k = \frac{2.(a + b.I_{mean})}{b.I_{mean}} . \tag{6.24}$$

If we know C_T, C_{target} and $C_{comparison}$ then we can calculate the value of k as

$$k = \left(\frac{|C_{target}^2 - C_{comparison}^2|}{C_T} \right) - (C_{target}^2 + C_{comparison}^2) , \tag{6.25}$$

and if we know the value k and the value of I_{mean} then we can subsequently calculate the ratio of b over a as

$$R = \frac{b}{a} = \frac{2}{I_{mean}.(k - 2)} . \tag{6.26}$$

The function $f(x) = ax + bx^2$ can be rewritten as $f(x) = p.(x + Rx^2)$ where p is a constant. Therefore the value R is a measure of the proportionate contribution of the quadratic component within the function $f(x)$.

When a contrast modulation is subject to a non-linearity, a first-order signal is introduced. In the contrast modulated noise stimuli employed in this chapter, the application of a non-

linearity will result in a signal that is very similar to a luminance modulation with additive noise. If we can find the detection threshold of a luminance modulated signal with additive noise that is similar in amplitude to the noise used in the experiments detailed in this study, then we can make an estimate of the value of C_T needed for detection of a contrast modulation. From the data described in the previous experiments, we already have data concerning the modulation depth and subsequent target and comparison contrasts necessary for detection of motion direction in contrast modulated noise. Therefore, if we can estimate C_T, we can then make an estimate of the value of R.

The experimental set up was the same as that for experiments 6.2 and 6.4. The stimulus was a luminance modulation with additive static standard binary noise. The mean luminance of the stimulus was 14.36 cd/m^2 and the amplitude of the binary additive noise was 4.31 cd/m^2, which gives a mean contrast of 0.3. The gamma correction procedure was the same as that employed in experiments 6.6 and 6.7. The single subject, CB, responded to 100 levels of prefixed luminance modulation contrast (0.25, 0.30, 0.35, 0.40 and 0.45%). The results are shown in Figure 6.17. The contrast detection threshold was calculated by probit regression as 0.36% (a contrast sensitivity threshold of 278). In experiment 6.4, the 75% threshold for subject CB with a target contrast of 0.3 was 0.104. This was calculated by probit regression from the data collapsed across targets and when comparison contrast was greater than target contrast (ie. at negative modulation depths). From this data, the value of k is 12.7 which gives a value for R of 0.013.

The calculation shows that the non-linearity required to produce effects of the kind measured in this study are fairly small. The small size of the necessary non-linearity would appear to argue against the notion that a specific second-order system is involved in the detection of motion direction in the second-order motion stimuli described in this chapter. If a second-order channel is involved then it must be somewhat ineffectual. It seems reasonable to propose that the results detailed in this study are based on small unintentional non-linearities within the visual system.

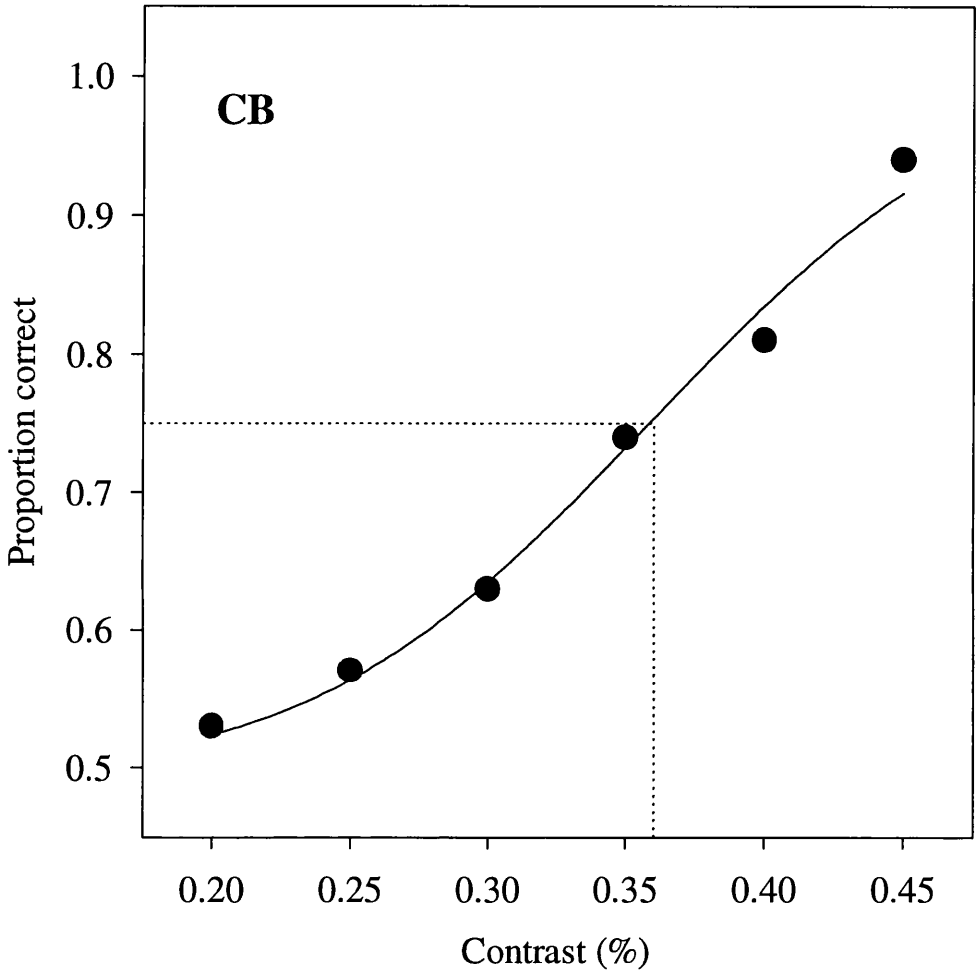


Figure 6.17: Results from experiment 6.8 showing proportion correct as a function of contrast for direction discrimination using a luminance defined motion sequence with additive noise. The black line shows a psychometric function fitted to the data using probit analysis. The dotted horizontal line marks the 75% correct level. The dotted vertical line marks the estimated contrast necessary for the subject to correctly judge motion direction on 75% of the trials.

6.3.5.2. Discussion of results from Experiment 6.8.

The data from experiment 6.8 shows that only a small non-linearity is necessary to account for modulation depth tuning curves for contrast defined motion. It might be argued that this conclusion is drawn from data for a single subject. However, the direction discrimination thresholds described in experiment 6.8 appear to be comparable to those obtained by other researchers. Figure 6.18 shows modulation depth detection thresholds for direction discrimination in luminance and contrast defined motion stimuli. These data are re-plotted from Nishida *et al* (1997). Included on the graph are the thresholds described in experiment 6.8. This Figure clearly demonstrates that the thresholds described in experiment 6.8 are in no way anomalous.

It would be interesting to obtain some measure to ascertain how much more effective the bandpass filtering plus rectification arrangement would be compared to the early non-linearity described above. The early non-linearity has been calculated to give a luminance contrast of 0.36% for a modulation depth of 0.10 with a target contrast of 0.3. This measure of luminance contrast has been taken to be equal to the contrast between the mean levels of the two noise patches after non-linear transformation. To gain some comparable figure for the second-order channel, we can extract the mean levels within the two patches of noise after bandpass filtering and rectification.

Space-time plots of motion sequences consisting of standard binary noise patches with a target contrast of 0.3, a modulation depth of −0.10 and a mean level of 14.36 were subject to bandpass filtering and rectification. The sequences contained static patterns and the width of noise patches was 32 times that of the noise (which had a width of 4 pixels). The stimulus was filtered with a Laplacian of Gaussian filter, the scale of the filter relative to the noise is shown in Figure 6.10. Each input image was 640 by 640 pixels. Results are drawn from a central region measuring 512 by 512 pixels. The output of the filtering and rectification processes appears much the same as a luminance defined version of the stimulus with additive bandpass noise. For each of 100 instantiations of the stimulus, the mean levels of the two constituent noise patches were measured and the contrast was calculated from these two values. The mean contrast was 9.59% with a standard deviation of 0.23%. Within each patch, the RMS contrast was also measured so that a measure of the noise present in the

image could be gained. The contrast of the noise (across patches and instantiations) was 65.91% with a standard deviation of 4.02%. Therefore, when bandpass filtering and rectification are applied to a contrast modulated noise stimulus, a first-order signal with a similar modulation depth to that of the original contrast modulation is produced. An additional effect is that bandpass additive noise is also present in the signal. The contrast of the additive noise appears to be approximately double that of the original noise carrier.

Unsurprisingly, the simulated second-order channel is doing a good job of extracting the contrast signal. If bandpass filtering and rectification were then followed by spatial filtering at the scale of the modulant (as suggested by Wilson *et al*, 1992), the bandpass noise present in the signal would largely be filtered out. If the subsequent motion energy detection stage was as effective as that proposed for the first-order channel, then thresholds for direction discrimination in contrast modulations should be similar to those for luminance modulations. However modulation depth detection thresholds for second-order stimuli appear to be approximately 25 times higher than those for first-order stimuli (Nishida *et al*, 1997). It is difficult to see how this difference can be reconciled with the notion of a second-order channel for the recovery of second-order motion.

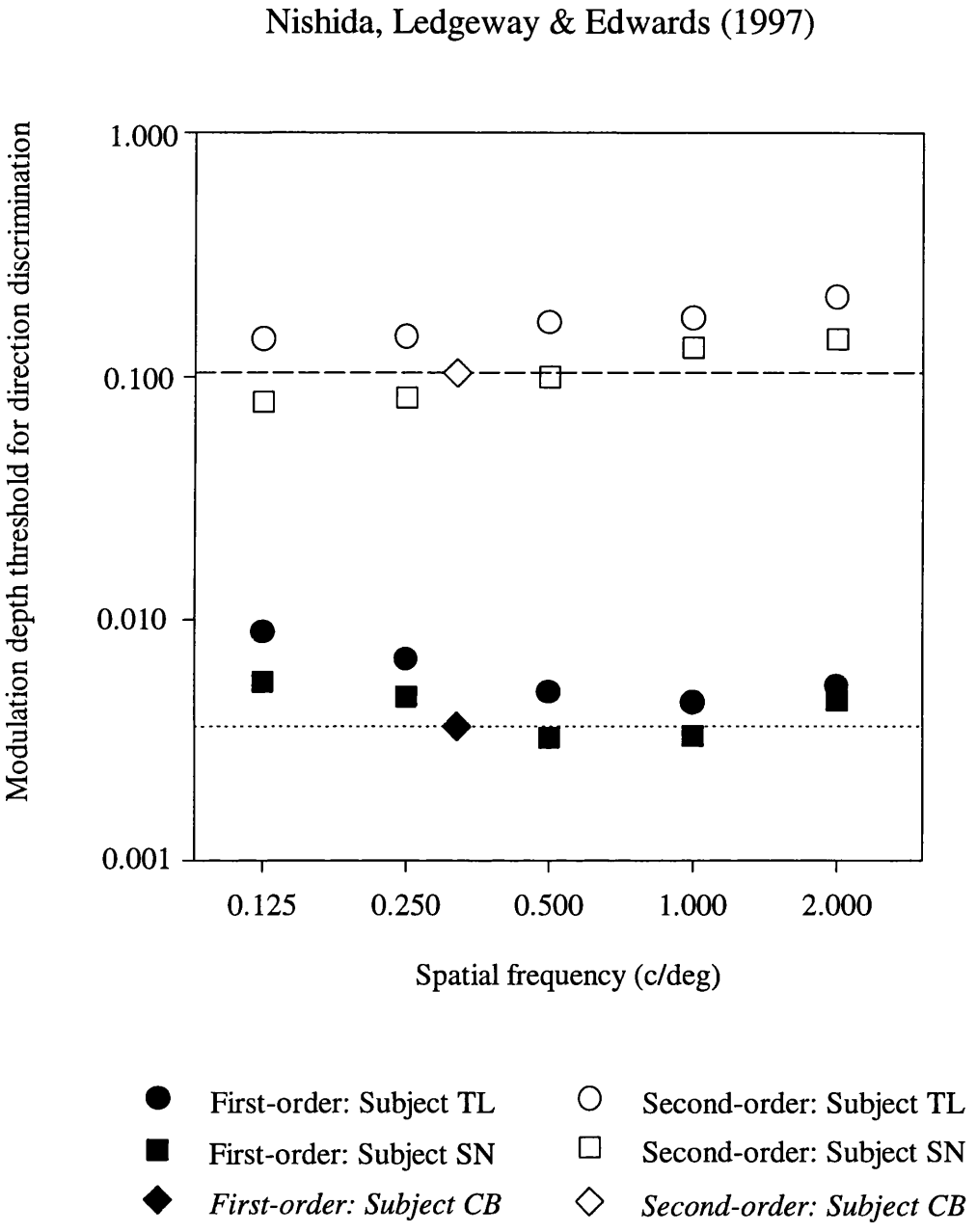


Figure 6.18: Data from Nishida, Ledgeway & Edwards (1997) (Figure 2, p2689) plotted with the modulation depth thresholds for subject CB described in experiment 6.8.

6.4. General Discussion.

One prominent finding that appears throughout this chapter is that a percept of motion is not elicited by stimuli containing motion defined by profound differences in luminance microstructure. In the characterisation provided by Cavanagh & Mather (1989) microstructure defined motion is a second-order stimulus, however the motion in this stimulus appears to be unavailable to the motion processing system. In itself, this presents an important constraint for any models that seek to account for the perception of second-order motion. Interestingly, when these patterns are static, it is relatively easy to discriminate between them. Therefore, with the stimulus parameters employed in this chapter, it appears that any active attentional feature tracking process, such as that described by Cavanagh (1991, 1992), does not play any part in the perception of motion in luminance microstructure defined motion stimuli.

At the simplest level, the results from these experiments show that the perception of contrast in motion and the perception of contrast itself are subject to some non-linearity. Both of these findings are consistent with a model in which the detection of contrast defined motion and the perception of contrast are dependent upon the output of bandpass spatial filters with some non-linear process occurring either at or after photodetection and prior to bandpass filtering. In terms of the detection of second-order motion, this favours the distortion product hypothesis over the notion of a specific second-order channel which is dependent upon some post-filtering non-linearity. The output of photodetectors is generally held to be adequately modelled by a compressive non-linearity (Tomita, 1968). Non-linearities in the cone response have been shown to exist by projecting interference fringes onto human retinae (MacLeod, Williams & Makous 1992; MacLeod & He 1993). When two interference fringes, each of which cannot be subjectively resolved, of different spatial frequencies or orientations are projected, then a beat pattern is seen. It may well be that these early non-linearities underlie the pattern of results described in this chapter.

There are two compelling pieces of evidence that appear to argue strongly against a role for early non-linearities. Firstly, the failure to null the distortion product by adding a luminance modulation to a second-order pattern (Badcock & Derrington, 1989) and secondly, the lack of coherent motion in stimuli consisting of interleaved contrast and luminance defined

motion (Ledgeway & Smith, 1994a; Mather & West, 1993). These findings may be accounted for by the idea that a range of both compressive and expansive early non-linearities are present in the human visual system. For example, we might suppose that, whilst the human visual system is broadly linear, attempts to correct for early compressive non-linearities might be imperfect. Whilst in some cases there may be a degree of undercompensation for the non-linearity, in other cases there may be a degree of overcompensation. In the example discussed here, where a compressive non-linearity is being corrected for, undercompensation will effectively result in a compressive non-linearity in the signal, whilst overcompensation will result in an expansive non-linearity in the signal.

When compressive and expansive early non-linearities are applied to a sinusoidal contrast modulation they will both give a signal at the frequency of the modulant. However, the two signals will be in counter-phase to one another. We might therefore think of the visual system as having lots of small non-linearities, the phase and amplitude of which might vary with space and time. If this is the case then the failure to null the distortion product by adding a luminance modulation becomes unsurprising. It is also possible that other parameters which may change the state of the visual system in its early stages may also affect non-linearities (ie. luminance and contrast).

In the interleaved contrast/luminance sequence described by Ledgeway & Smith (1994a), a modulant translates forward by a quarter of a cycle every frame and is used to alternately modulate luminance and contrast. Therefore, if one frame contains a luminance defined grating, the next frame will contain a contrast defined grating. When a compressive non-linearity is applied to a contrast modulation, a first-order component of the same frequency but the opposite phase to the modulant is introduced into the signal. In the case of an expansive non-linearity, the first-order component will have the both same frequency *and* phase as the modulant. In the interleaved contrast/luminance stimulus, any motion detector operating on part of the image affected by an expansive non-linearity will signal motion in the direction of modulant translation. However, any motion detector operating on areas of the image affected by a compressive non-linearity will signal motion in the reverse direction. We might therefore expect motion in opposite directions to be present in the image but there may be little or no chance of any particular direction becoming dominant.

By extending the distortion product hypothesis a number of results that appear to argue against the hypothesis can be explained. In this reformulation of the hypothesis, it is necessary to propose that both compressive and expansive early non-linearities exist in the visual system. This in turn can be accounted for by supposing that some attempt is made to correct for non-linearities inherent in the system but that this process is imperfect and that undercompensation and overcompensation for these non-linearities leads to the introduction of various expansive and compressive non-linearities.

Early non-linearities in the second-order channel?

It is possible to argue that the results detailed in this chapter do not necessarily favour the distortion product hypothesis over the notion of a second-order channel. One might reasonably propose that both pre- and post-filtering non-linearities may operate in the second-order channel. Because the second-order channel is non-linear anyway, a few pre-filtering non-linearities might be relatively unimportant. There is no real reason why the system should strive for linearity when it is fundamentally non-linear. Such a mechanism might well show the pattern of results outlined in this study.

It appears however that the second-order system is potentially far too effective at recovering the modulant to account for the width of the U-shaped modulation depth tuning curves. An estimate of the required size of the non-linearity shows that only a small non-linearity is necessary to account for the data. Surely a second-order channel designed specifically to extract contrast motion would display a higher level of performance? An additional argument against the notion of early non-linearities in the second order channel may be based on the tuning of the filters in the channel. The operation of an early non-linearity upon the input signal will mean that, after the non-linear transformation, the signal is essentially a contrast modulation with some small added luminance modulation at the frequency of the modulant. Wilson, Ferrara & Yo (1992) and Sutter, Sperling & Chubb (1995) suggest that pre-rectification filters in the second-order channel are tuned to higher spatial frequencies than the post-rectification filters. If the tuning of these two sets of filters differ in any significant way it is difficult to see how any first-order modulation in the signal can reasonably affect the detection of second-order motion as the first-order signal should be drastically attenuated by the filtering processes.

Fullwave stimuli.

The distortion product hypothesis cannot account for the perception of forward motion elicited by the reversed-phi motion sequence described by Chubb & Sperling (1989a). Chubb & Sperling proposed that the forward motion is based upon a fullwave rectification. Variations of the stimulus have also been used as proof for a fullwave-like process (Solomon & Sperling, 1994; Sperling, Chubb, Solomon & Lu, 1994). Figure 6.19a shows a space-time plot of the reversed-phi stimulus and two Fourier spectra. The topmost Fourier spectrum is from the reversed-phi stimulus, the bottommost spectrum is from a version of the stimulus where the value of the luminance has been squared. As luminance is always positive, this is not a squaring around mean luminance, but represents the quadratic distortion of an early non-linearity. Although this squaring process does introduce components that would indicate motion in the forward direction, these are of the same spatial and temporal frequencies but of lower amplitude than components indicating motion in the reversed direction. Therefore early non-linearities cannot readily account for the percept of forwards motion in this stimulus.

The fact that one example of second-order motion might be detected through the operation of an early non-linearity, coupled with the fact that the forward motion in the reversed phi stimulus cannot be explained by an early non-linearity, implies that either these two stimuli are processed in different manners under the auspices of a single system, or are processed by separate mechanisms. This conclusion is strengthened by results from a number of other researchers who have proposed that second-order motion may be detected through an early non-linearity and also through the operation of a non-Fourier mechanism.

Holliday & Anderson (1994) examined the detection of motion direction in beat patterns and sine wave gratings over modulant temporal frequency. The modulation depth sensitivity function for the beat pattern was bimodal, whilst that for the luminance modulation was unimodal. In an additional experiment which examined cross adaptation between the two types of stimulus, Holliday & Anderson showed that adaptation tuning curves were similar for the two stimulus types when the temporal frequency of the beat modulation was high. This was not the case when the beat modulation temporal frequency was low (below about 4 Hz). Holliday & Anderson argued that these results show that second-order motion is

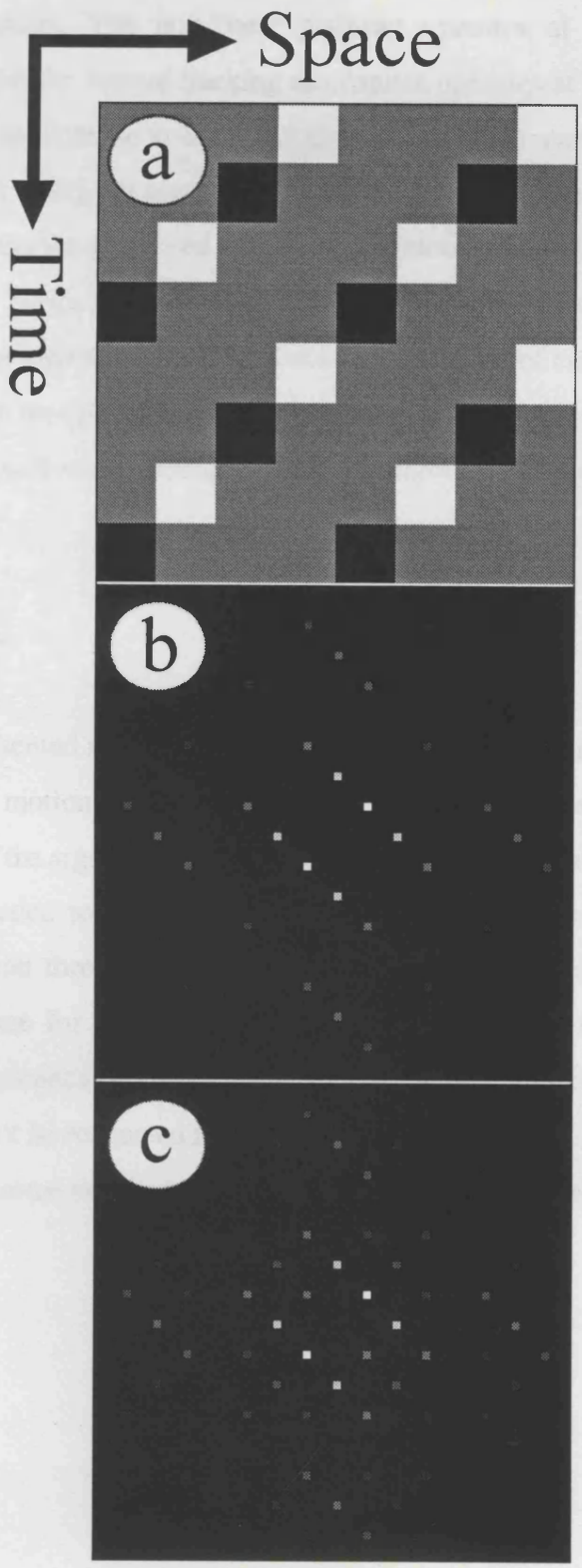


Figure 6.19: (a) Space-time plot of Chubb & Sperling's (1989a) reversed-phi motion stimulus, (b) Fourier spectrum of stimulus, and (c) Fourier spectrum of the stimulus when subject to squaring non-linearity applied directly to luminance.

detected by two mechanisms. They identify one mechanism as first-order motion energy detection operating on a distortion product. The other mechanism is identified as a feature tracking mechanism. The non-linear pathway operates at high modulant temporal frequencies whilst the feature tracking mechanism operates at low temporal frequencies. Using a nulling paradigm, Scott-Samuel & Georgeson (1995) reach a similar conclusion (see also Scott-Samuel, 1997). In contrast to Badcock & Derrington (1989), they managed to successfully cancel the perceived motion of a contrast modulation by adding a luminance signal, but only when the modulant temporal frequency was high. Scott-Samuel (1997) proposed that at high modulant frequencies, the motion of contrast modulation may be perceived through the operation of an early non-linearity but that at low modulant temporal frequencies, second-order motion is perceived through the operation of a second-order channel.

6.5. Conclusion.

The evidence presented in this chapter favours the notion that distortion products underlie the detection of motion direction of the second-order motion stimuli investigated in this chapter. Some of the argument that supports this conclusion rests upon the small size of the non-linearity needed to account for the data. This in turn is based upon measures of direction detection thresholds in contrast modulated and luminance modulated stimuli. Stronger evidence for a role for distortion products comes from a discrepancy when comparing performance on stimuli composed of top biased and bottom biased noise. This discrepancy cannot be accounted for by an approach based on initial linear bandpass filtering and implies that some non-linearity exists prior to bandpass filtering.

Chapter Seven

Discussion.

7. Discussion.

The central motivation behind the simulations and experiments described in this thesis is to examine the hypothesis that motion is extracted from the analysis of a Fourier decomposition of the signal. More specifically the hypothesis states that motion is present as orientations of Fourier energy through the frequency domain origin. The existence of second-order motion stimuli has challenged this hypothesis and it has been proposed that some additional mechanism exists that deliberately extracts second-order motion. A crucial stimulus in this debate is the reversed-phi motion sequence described by Chubb & Sperling (1989a). In terms of a non-linearity applied to the signal, the forward motion in this stimulus can only be made available to a standard motion energy analysis through the operation of a fullwave rectification-like non-linearity. Whilst it is possible that some second-order motion may be perceived through the operation of a monotonic non-linearity applied to the luminance signal (the distortion product hypothesis), the perception of forward motion in the reversed-phi stimulus shows that some other mechanism must operate. The important point to be considered here is whether this mechanism operates independently of a first-order mechanism or whether both first- and second-order motion are perceived through the operation of some single mechanism. If both types of motion are perceived through the operation of a single system then the hypothesis that motion is extracted as orientations of Fourier energy through the frequency domain origin must be revised or abandoned. Alternatively, in order for the hypothesis to be supported, one must suppose that some non-Fourier mechanism exists. Therefore the issue of whether first- and second-order motion are perceived by the same or different mechanisms has a critical bearing on the question of how motion is perceived generally.

7.1. The carrier is crucial.

A number of studies have sought to provide evidence for or against a difference in processing between first- and second-order motion. In general these have not isolated any substantial perceptual differences between the two types of stimulus, and many appear to provide evidence for similarities in response to first- and second-order motion (see introduction for a brief review). Rather than abandon the hypothesis that first- and second-order motion are perceived through the operation of different mechanisms, one approach has been to suggest

that the similarities in response merely reflect similarities in the operation of the two separate mechanisms. The similarities are explained by supposing that first- and second-order motion are extracted through motion energy analysis but that some demodulative process is applied to the second-order signal prior to motion energy extraction. Within this context it is difficult to see how one might reasonably provide evidence supporting the hypothesis that first- and second-order motion are perceived through the same mechanism. Surely, a similarity in response to the two different stimulus types can also be seen as evidence for the hypothesis that both first- and second-order motion are perceived through the operation of a single mechanism?

We might go a step further by supposing that differences in response to two different stimuli do not necessarily reflect the operation of two different mechanisms. In Chapter 3 a series of simulations showed that a standard motion energy mechanism would not indicate coherent motion to contrast modulations of both static and dynamic noise. This work was motivated by a finding described by Smith & Ledgeway (1997a) which showed that, for contrast modulations of dynamic noise, direction discrimination thresholds were considerably higher than orientation discrimination thresholds. With a static noise carrier the two thresholds were similar. Smith & Ledgeway proposed that stochastic biases in the underlying noise carrier introduced first-order components and this effect was more pronounced in the case of the static noise carrier than in the case of the dynamic noise carrier. Based upon this hypothesis they proposed that the modulations of static noise were detected by a first-order mechanism whilst the higher direction discrimination thresholds found in response to the contrast modulated dynamic noise reflected the operation of a second-order system.

The data from the simulations described in Chapter 3 indicate that the analysis presented by Smith & Ledgeway (1997a) is incorrect. Contrast modulations of both static and dynamic noise must be detected through some process other than standard motion energy analysis. Unless we wish to propose that these stimuli are detected through two different mechanisms, the difference in response shown to the two stimulus types may well reflect some simple interaction between the stimuli and the nature of the psychophysical task. If one analyses the Fourier energy content of the stimuli, then it appears that we have a good reason for proposing that motion direction detection thresholds for contrast modulated

dynamic noise should be higher than those found for contrast modulated static noise. In dynamic noise stimuli there is simply more random "motion noise" than in the static noise stimuli. If one is making a judgement on motion direction, the dynamic noise stimulus is simply noisier in relation to that particular task than the static noise stimulus. It therefore appears that we have here an example of an instance where a difference in response does not necessarily imply different mechanisms.

The results of a series of experiments examining speed discrimination thresholds (SDTs) in first- and second-order bars and edges are described in Chapter 4. Both of the second order stimuli consisted of contrast modulations of binary noise. Whilst there were no consistent differences in the patterns of response over velocity and stimulus width, SDTs for the second-order stimuli appeared to be slightly higher than those for the first-order stimuli. This difference was more pronounced for the bar stimuli than for the edge stimuli. This latter finding was accounted for by proposing that there was a higher proportion of motion noise in the second-order bar stimulus than in the second-order edge stimulus. The difference between the first- and second-order stimuli can also be explained by appealing to the same mechanism: the second-order stimuli contain more random motion noise than the first-order stimuli. It is by no means unlikely that random fluctuations in the local velocity fields of the second-order stimuli should affect SDTs. Again, it appears that the difference between first- and second-order SDTs may be explained purely with reference to task specific noise introduced by the carrier, rather than by appealing to some difference in processing.

In Chapter 5, the perceived direction of motion in Chubb & Sperling's (1989a) reversed-phi stimulus was examined over a number of stimulus parameters. Additionally, the perception of motion direction in a contrast defined version of this stimulus was investigated. Both luminance and contrast defined reversed-phi motion sequences showed similar patterns of response over spatial and temporal frequency. Where differences did occur, this could be put down to an effect of the noise carrier in the contrast defined stimulus. This was confirmed by adding noise to the luminance defined stimulus and showing that the pattern of response became similar to that of the contrast defined stimulus. It was also shown that the type of carrier could have a large effect on the perception of motion direction in the contrast defined stimulus. When a static carrier was replaced by a dynamic carrier, the perceived direction of motion could be completely reversed. The evidence therefore

suggests that the choice of carrier in second-order motion stimuli is crucial to the performance of subjects on psychophysical tasks.

7.2. A second-order channel?

In terms of the standard two channel approach, the fact that both forward and reversed motion can be elicited by both luminance and contrast defined reversed-phi stimuli necessitates the extension of this model to include some third channel that can account for the existence of the forward motion percept elicited by the contrast defined stimulus. However, the fact that the patterns of response for both types of stimulus are so similar must provide some support for the hypothesis that both first- and second-order motion are extracted through the same mechanism.

Some of the similarities and cross-over effects between first- and second-order motion might adequately be explained if the motion of some second-order stimuli were perceived by virtue of some small distortion product in a largely "linear" channel. In this thesis (see Chapter 6) the distortion product was reformulated in the following manner. Assume that the human motion system is broadly linear but that there are small accidental non-linearities present in the system. This simply reflects the fact that perfect linearity in a analogue biological system is probably somewhat difficult to obtain. If this is the case then there should be no expectation that these non-linearities can be characterised by any single non-linearity applied to the input pattern. At different points in both space and time the input image may be subject to either compressive or expansive non-linearities. If the visual system does attempt to be linear then this reformulated distortion product hypothesis may well provide a better description of the operation of distortion products within the "linear" channel. If there was some single known non-linearity applied to the input image then it would seem possible that the visual system could correct for the operation of this non-linearity. A technique for the correction of quadratic non-linearities is described by Derrington (1990) who noted that, although there appear to be non-linearities in the X-cells of the lateral geniculate nucleus (Derrington, 1987), the response of simple cells in the striate cortex appears to be linear (Albrecht & DeValois, 1981).

The reformulation of the distortion product hypothesis in terms of a distortion distribution is useful because it allows us to account for some of the data that previously appeared to provide strong evidence against a role for distortion products. If both expansive and compressive non-linearities are present then an attempt to null one particular type of non-linearity should be unsuccessful (e.g. Badcock & Derrington, 1989). Smith & Ledgeway (1994a) presented subjects with an interleaved first- and second-order stimulus. In this, luminance and contrast modulations are temporally interleaved with a ¼ cycle spatial phase shift between each subsequent modulant (for both pattern types the modulant has the same spatial frequency). An example of an interleaved luminance/contrast modulation stimulus is shown in Figure 7.1a. The fact that motion is not readily perceived in these stimuli (at least at low temporal frequencies) has been taken as positive evidence for the existence of separate first- and second-order channels.

If the interleaved luminance/contrast stimulus was subject to an expansive non-linearity, motion should be seen in the direction of modulant motion. This is because the non-linearity would introduce a first-order component of the same phase and frequency as the modulant when applied to the contrast modulation. However, a compressive non-linearity would introduce a component that is half a cycle out of phase with the modulant and would therefore result in reversed motion. If the non-linearities are balanced, then no one direction should predominate. This latter point provides a simple counter to the assertion that the lack of coherent motion elicited by this stimulus proves the existence of separate first- and second-order channels.

The experiments described in Chapter 6 of this thesis provide evidence for a distortion product in the perception of motion in contrast modulated noise stimuli. It is not clear what role is played by these distortion products. The results also show that there is some degree of non-linearity in the perceived contrast of static noise patterns. Scott-Samuel & Georgeson (1995) provided evidence which appeared to show that distortion products fail to operate at low modulant temporal frequencies. Based on their data, there should be no strong distortion product operating at the modulant temporal frequency employed in the experiments in chapter 6. However it should be noted that the techniques employed by Scott-Samuel & Georgeson would fail to provide evidence for a non-linearity if the magnitudes of spatio-temporally localised expansive and compressive non-linearities were

globally balanced. In other words, the tests employed are only sensitive if one type of non-linearity predominates. This means that the effect that was measured may well reflect (or partially reflect) the increasing dominance of one type of non-linearity rather than the absolute magnitudes of the non-linearities in general.

The evidence would therefore appear to show that we cannot discount the distortion product hypothesis. This in turn means that a number of the results in the second-order literature and in this thesis may possibly be based upon such non-linearities. However it should be emphasised that the distortion product hypothesis, in combination with the oriented Fourier energy hypothesis, cannot account for the perception of forward motion in both luminance and contrast defined reversed-phi motion sequences. Additionally, the similarities in the patterns of response to both the luminance and contrast defined motion sequences appear to indicate that both stimulus types are processed by the same mechanism. If it were supposed that some single linear mechanism could account for the perception of both forward and reversed motion in the luminance defined reversed-phi stimulus, then the application of a distortion product prior to the operation of this mechanism could potentially account for the similarity in response shown by the luminance and contrast defined stimuli.

In summary, on the basis of the experiments conducted in this thesis, there is no convincing support for the hypothesis that a separate second-order channel exists in addition to a first-order channel. The evidence appears to support the hypothesis that both first- and second-order motion are perceived by a single mechanism. The picture is complicated by the finding that distortion products may play some part in the perception of second-order motion. However distortion products cannot account for the perception of forward motion in both the luminance and contrast defined reversed-phi stimuli, at least in the context of the oriented Fourier energy approach. There are potentially two theories that can reasonably account for the similarity in response to these two stimuli. Firstly, that some single mechanism mediates both the perception of forward and reversed motion in the luminance defined stimulus but that some small distortion product makes the modulant in the contrast defined stimulus available for motion analysis. Secondly, that some mechanism mediates the perception of both forward and reversed motion in both the luminance and contrast defined reversed-phi stimuli without recourse to some specific demodulative process. Both of these

options imply that some degree of change is necessary for the oriented Fourier energy hypothesis.

7.3. Alternative Models for the extraction of second-order motion.

The findings presented in this thesis support the hypothesis that some second-order motion is perceived through the same mechanism by which first-order motion is perceived. At least two extant models can account for some instances of both first- and second-order motion. One such approach is an extended gradient model (Johnston *et al*, 1992; McOwan & Johnston, 1995). This model has been shown to account for the perception of motion in contrast modulated sine wave gratings (Johnston & Clifford, 1995a) and has also been shown to exhibit a direction reversal with increases in the spatial frequency of the luminance defined reversed-phi stimulus (Johnston & Clifford, 1995b). If some early non-linearity were applied to the contrast defined version of this stimulus, then it is possible that the model would also show such an effect with this stimulus. The model cannot account for the motion reversal with changes in the temporal frequency of the luminance defined stimulus. A recent extension of this model from one spatial dimension to two spatial dimensions has been shown to detect the motion of contrast modulations of static and dynamic noise without recourse to some non-linearity prior to motion extraction (Johnston & Benton, 1996). Although this model can extract both first- and second-order motion, it shows no directional preference when applied to Ledgeway & Smith's (1994a) interleaved first- and second-order stimulus (Johnston & McOwan, 1996).

Another model that can account for both first- and second-order motion is an extension of a feature model (Watt & Morgan, 1985) into the motion domain (Morgan, personal communication). In this, the halfwave rectified outputs from the on-channel are combined into one processing stream whilst the halfwave rectified outputs of the off-channel are combined into another processing stream. Each processing stream is the sum of the outputs from filters tuned to a range of spatial frequencies. A procedure to identify zero bounded regions is then applied to each processing stream. Motion direction is calculated on a nearest zero bounded region basis. With relation to the reversed-phi stimulus described by Chubb & Sperling (1989a), the model will increasingly indicate forward motion with increases in the spatial scale of the stimulus (see Chapter 5). The argument is based upon the properties

Figure 7.1: (a) Space-time plot showing an interleaved luminance/contrast modulation stimulus. The results of halfwave rectification around mean luminance are shown in (b) and (c) with: (b) values above mean luminance and (c) values below mean luminance.

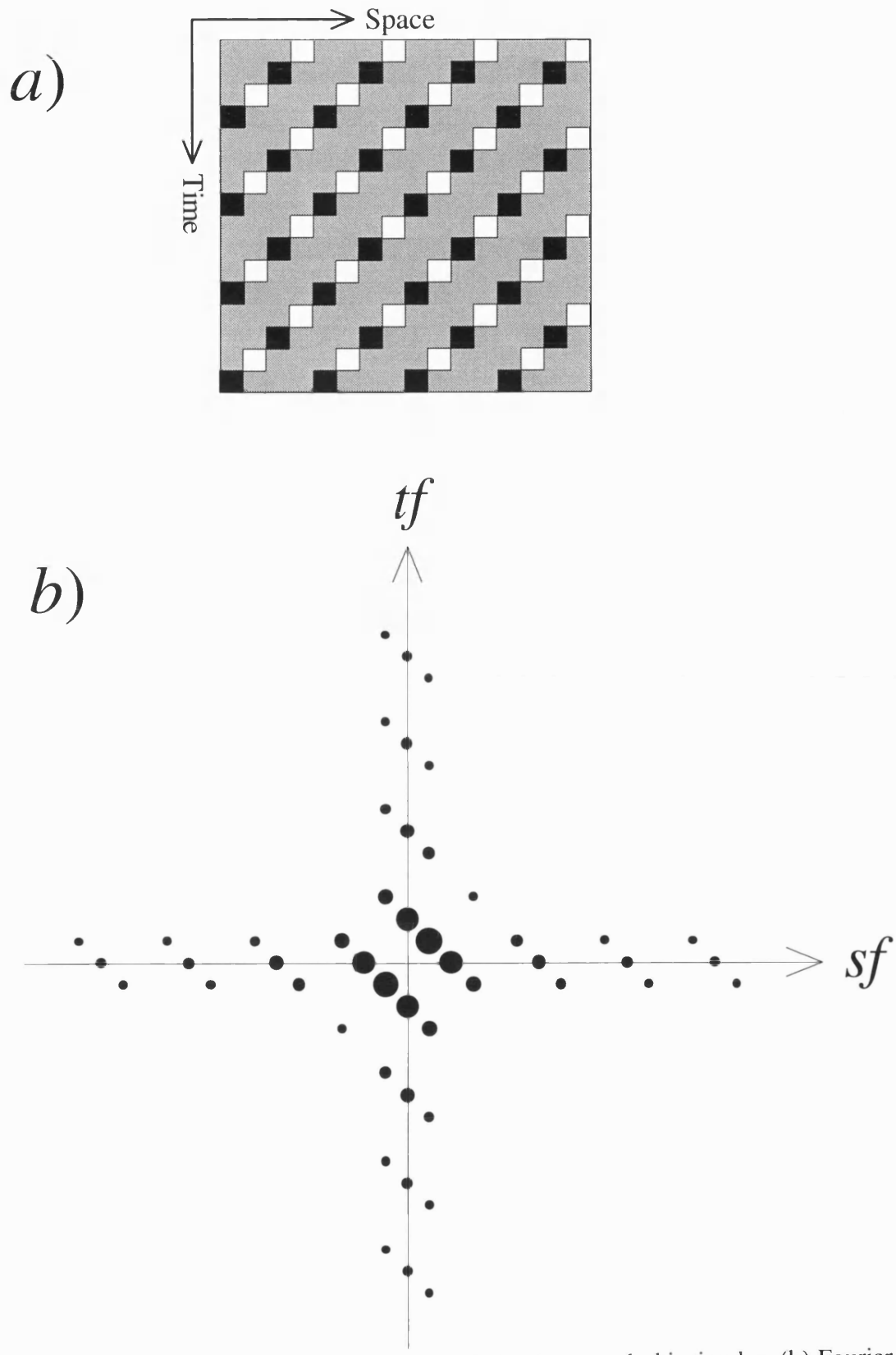


Figure 7.2: (a) Space-time plot of a luminance defined reversed-phi stimulus, (b) Fourier spectrum of the stimulus. The area of the dots indicating each component is directly proportional to the amplitude of the component. Those components with an amplitude less than one tenth of the largest component are not shown.

of the spatial filters in relation to the scale of the stimulus. The model would have to be explicitly extended to cover temporal filtering if one wished to explain the reversal of motion direction for the reversed-phi stimulus over temporal frequency. The model can also account for the lack of consistently perceived motion in the interleaved luminance modulation/contrast modulation stimulus. The on-channel stream will signal motion in one direction whilst the off-channel stream will signal motion in the opposite direction. This is shown in Figure 7.1 where the results of applying halfwave rectification to an interleaved luminance/contrast stimulus are displayed.

An acceptance of the models discussed above implies a departure from the oriented Fourier energy hypothesis. One particular strength of the standard Fourier energy hypothesis is that the operation of the mechanism can be predicted on a simple intuitive base, this is not the case with the models described above. It might be well worth considering what the minimal paradigm shift may be in order to account for both first- and second-order motion without recourse to a non-Fourier channel. As noted by Fleet & Langley (1994), the motion of many second-order stimuli can be characterised by orientations of Fourier motion energy that do not lie on lines or planes containing the frequency domain origin. This observation does not simply apply to stimuli such as translating beats or contrast modulations but also to more complex stimuli such as, for example, the reversed-phi stimulus of Chubb & Sperling (1989a). Figure 7.2 shows the Fourier spectrum of this stimulus. The areas of the dots representing the components are directly proportional to the amplitude of the components. We can readily see that local second-order orientations of Fourier energy are present in the Fourier spectrum and that these signal motion in the direction opposite to that signalled by the largest component (ie. the component signalling reversed motion).

One "minimum paradigm shift" hypothesis is, therefore, that motion may be extracted from oriented Fourier components, but that the orientation of these components does not necessarily have to be through the frequency domain origin. If this were the case then two types of process might be applied to an array of motion energy detectors. In one, the orientations of Fourier energy through the frequency domain origin are recovered to give a measure of first-order velocity. In the other, local (in terms of frequency space) orientations of Fourier motion energy are measured to give a measure of second-order velocity. In this hypothesis, both first- and second-order motion are recovered by different

processes but are dependent upon the same array of motion energy detectors. If this scheme could be implemented, then this might be more effective (in terms of system resources) than having a separate array of motion energy detectors within some second-order channel. In addition one might propose that small non-linearities inherent in the system also operate so that, in some cases, second-order motion can be detected through the process recovering orientations of Fourier energy through the frequency domain origin.

It is not however clear how such a scheme could be implemented. The majority of approaches to the problem of extracting velocity from arrays of motion energy detectors primarily use some combination of outputs from detectors tuned to the same spatial frequency (e.g. Adelson & Bergen, 1985). By using this approach the temporal frequency can be estimated. If this is done within a number of spatial frequency channels and if some confidence measure is attached to each measurement, then the velocity can be calculated from the output of the spatial frequency channel with the largest associated confidence measure. However, to recover orientations of Fourier motion energy that do not lie through the frequency domain origin it is necessary to establish some estimate of the spatio-temporal relationship between the components. Therefore a scheme for the recovery of second-order Fourier energy orientations must involve some combination of outputs from motion energy detectors tuned to different temporal and different spatial frequencies.

A simple scheme for the recovery of second-order motion direction may be implemented as follows. Let us consider the opponent energy arrangement described by Adelson & Bergen (1985). This consists of a pair of motion energy detectors tuned to the same spatial frequency and the same absolute temporal frequency but maximally responsive to motion in opposite directions. If we take two pairs of Adelson & Bergen opponent energy detectors (*A* and *B*) that are tuned to different spatial frequencies then we can have the arrangement, in terms of frequency space, shown in Figure 7.3. By combining these in the following manner:

$$Output = A_L.B_R - A_R.B_L$$

we have a mechanism that can detect the direction of motion in contrast modulations and beats. The model will give a maximum positive output when A_L and B_R are stimulated simultaneously and will give a maximum negative output when A_R and B_L are stimulated

simultaneously. The response of the model to some second-order patterns is shown in Figure 7.4. This model is too flawed to be thought of as a mechanism for the extraction of second-order motion in humans. In particular, it is difficult to see how any measure of velocity (rather than direction) could be calculated. Furthermore, the model will respond to first-order motion if the spatio-temporal tuning curves of A_L and B_R or A_R and B_L overlap.

The mechanism described above is offered for the purpose of illustration only. It shows that second-order motion can be extracted from combining the outputs of motion energy detectors tuned to different spatial and temporal frequencies. Additionally, the model demonstrates that second-order motion can be extracted without the use of some non-linearity to demodulate the signal.

In summary, there are a number of models that can potentially account for some first- and second-order motion without recourse to a separate mechanism. It is possible that the impact of these models has been weakened by the fact that the results of their operation upon a given image are not immediately obvious. This does not of course mean that they incorrect, merely that they are inconvenient. Alternatively, the hypothesis that the human visual system is capable of detecting orientations of Fourier energy that are not oriented through the frequency domain origin has been proposed. Such a theory has the advantage that it requires a comparatively limited revision of the current viewpoint but it is by no means clear how such a mechanism could be implemented.

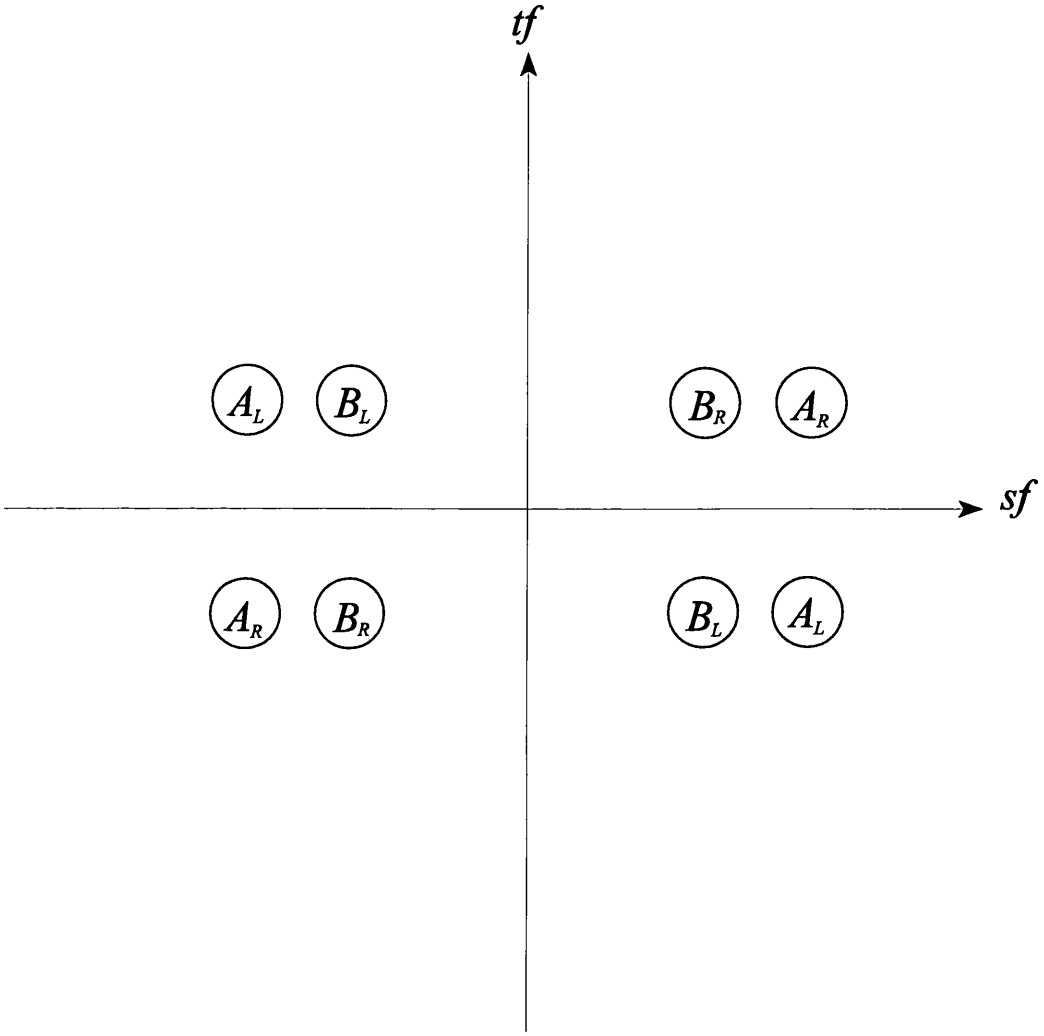


Figure 7.3: Position of two opponent energy pairs *A* and *B* of motion energy detectors in frequency space. A subscripted *R* indicates that the detector preferentially responds to rightward motion, a subscripted *L* indicates that the detector preferentially responds to leftward motion.

Figure 7.4: (*on following page*) Results of model (see text for details) when applied to space-time plots of: (a) beat translating to the left, (b) beat translating to the right, (c) contrast modulation of static sine wave grating translating to the left, (d) contrast modulation translating to the right and (e) contrast modulation with sinusoidally modulated motion. Input patterns are shown on the left hand side, output patterns on the right hand side. In the output patterns, areas lighter than the border indicate leftward motion whilst areas darker than the border indicate rightward motion.

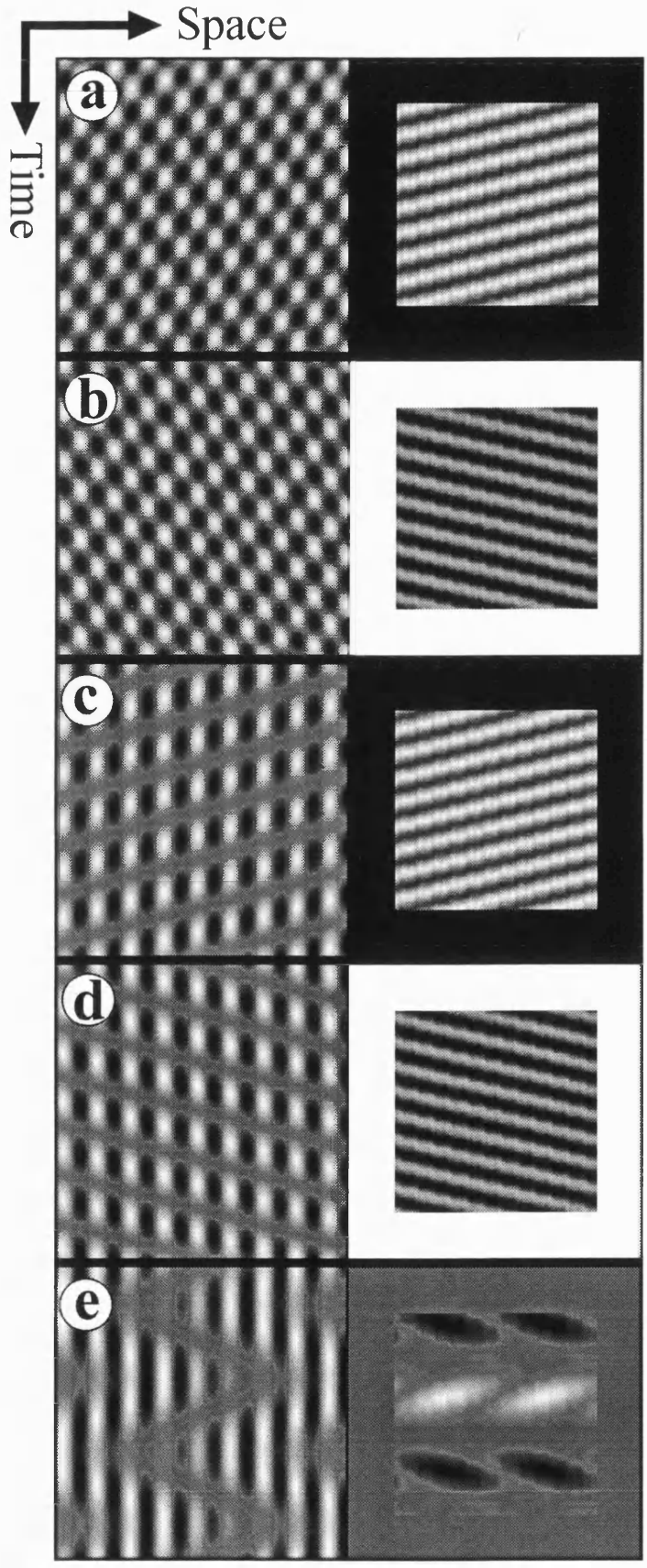


Figure 7.4: *Caption on previous page.*

7.4. Conclusions.

The issue of whether first- and second-order motion are perceived through the operation of separate channels has a critical bearing upon on current theories of the perception of motion. Evidence for the operation of a separate channel can come from data showing some difference in response to the two stimulus types. The analysis presented in this thesis implies that differences in performance do not necessarily imply different mechanisms. Dissimilarities in response to first- and second-order motion may well be put down to the presence of the carrier rather than the operation of separate first- and second-order channels. It may be possible to check whether this is the case by adding the carrier to a first-order signal to see whether these differences are eradicated. In the context of the two channel model, findings presented in this thesis necessitate the addition of a third channel to account for the data. A simpler interpretation is that first- and second-order motion may be detected through the operation of a single mechanism. Additional experiments have shown that distortion products may well play a part in the perception of second-order motion. It is noted that there are at least two extant models which can account for some aspects of both first- and second-order motion without recourse to a separate second-order channel. An additional hypothesis is proposed in which the orientations of Fourier motion energy that do not fall through the frequency domain origin are recovered. It is not yet clear whether such a scheme could adequately be implemented.

Chapter Eight

References.

8. References.

Adelson, E.H. & Bergen, J.R. (1985). Spatio-temporal energy models for the perception of apparent motion. *Journal of the Optical Society of America A*, 2, 284-299.

Adelson, E.H. & Bergen, J.R. (1986). The extraction of spatio-temporal energy in human and machine vision. *In Motion: Representation and Analysis (IEEE Workshop Proceedings)*, pp. 151-155.

Albrecht, D. & DeValois, R.L. (1981). Striate cortex responses to periodic patterns with and without fundamental harmonics. *Journal of Physiology, London*, 319, 497-514.

Albright, T.D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, 52, 1106-1130.

Albright, T.D. (1992). Form-cue invariant motion processing in the primate visual cortex. *Science*, 255, 1141-1143.

Anderson, S.J., Drasdo, N. & Thompson, C.M. (1995). Parvocellular neurons limit motion acuity in human peripheral vision. *Proceedings of the Royal Society of London B*, 261, 129-138.

Anderson, S.J., Holliday, E.I., Singh, K.D. & Harding, G.F.A. (1996). Localization and functional analysis of human cortical area V5 using magneto-encephalography. *Proceedings of the Royal Society of London B*, 263, 423-431.

Anstis, S.M. (1970). Phi movement as a subtraction process. *Vision Research*, 10, 1411-1430.

Anstis, S.M. & Rogers, B.J. (1986). Illusory continuous motion from oscillating positive-negative patterns: implications for motion perception. *Perception*, 15, 627-640.

Badcock, D.R. & Derrington, A.M. (1989). Detecting the displacement of spatial beats: No role for distortion products. *Vision Research*, 29, 731-739.

Baker, C.L., Hess, R.F. & Zihl, J. (1991). Residual motion perception in a "motion-blind" patient, assessed with limited-lifetime random dot stimuli. *Journal of Neuroscience*, 11, 454-461.

Beckers, G. & Zeki, S. (1995). The consequences of inactivating area V1 and V5 on visual motion perception. *Brain*, 118, 49-60.

Benton, C.P., Johnston, A.J. & McOwan, P.W. (1996). Testing the rectification hypothesis for second-order motion perception. *Investigative Opthamology & Visual Science (Suppl.)*, 37, 746.

Bowne, S.F., McKee, S.P. & Glaser, D.A. (1989). Motion interference in speed discrimination. *Journal of the Optical Society of America A*, 6, 1112-1121.

Bowns, L. (1996). Evidence for a feature tracking explanation of why type II plaids move in the vector sum direction at short durations. *Vision Research*, 36, 3685-3694.

Braddick, O.J. (1974). A short-range process in apparent motion. *Vision Research*, 14, 519-527.

Braddick, O.J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society, London B*, 290, 137-151.

Braunstein, M.L. (1994). Structure from motion. In Smith, A.T. & Snowden, R.J. (Eds), *Visual Detection of Motion.* London: Academic Press.

Brown, J. & Glazier, E.V.D. (1980). *Telecommunications.* London: Chapman and Hall.

Bruce, V., Green, P.R. & Georgeson, M.A. (1996). *Visual Perception.* Hove: Psychology Press.

Burbeck, C.A. (1981). Criterion-free pattern and flicker thresholds. *Journal of the Optical Society of America*, 71, 1343-1350.

Burbeck, C.A. (1987). Position and spatial frequency in large-scale localisation judgments *Vision Research*, 27, 417-427.

Burton, G.J. (1973). Evidence for non-linear response processes in the visual system from measurements on the thresholds of spatial beat frequencies. *Vision Research*, 13, 1211-1225.

Campbell, F.W. & Robson, J.G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, London*, 197, 551-566.

Carney, T. (1995). Cyclopean motion: Feature tracking or motion energy. *Investigative Opthamology & Visual Science (Suppl.)*, 36, 52.

Carney, T. (1997), Evidence for an early motion system which integrates information from the two eyes. *Vision Research*, 37, 2361-2368.

Carney, T. & Shadlen, M.N. (1992). Binocularity of early motion mechanisms: comments on Georgeson and Shackleton. *Vision Research*, 32, 187-191.

Carney, T. & Shadlen, M.N. (1993). Dichoptic activation of the early motion system. *Vision Research*, 33, 1977-1995.

Cavanagh, P. (1991). Short-range vs long-range motion: not a valid distinction. *Spatial Vision*, 5, 303-309.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257, 1563-1565.

Cavanagh, P. & Mather, G. (1989). Motion: the long and the short of it. *Spatial Vision*, 4, 103-129.

Celesia, G.G. & DeMarco, P.J., Jr. (1994). Anatomy and physiology of the visual system. *Journal of Clinical Neurophysiology*, 11, 482-492.

Chubb, C., McGowan, J., Sperling, G. & Werkhoven, P. (1994). Non-Fourier motion analysis. In Bock, G.R. & Goode, J.A. (Eds), *Higher order processing in the visual system: Ciba Foundation Symposium, 184.* (pp.193-210). Chichester: Wiley.

Chubb, C. & Sperling, G. (1988). Drift-balanced random dot stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A,* 5, 1986-2007.

Chubb, C. & Sperling, G. (1989a). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences, U.S.A.,* 86, 2985-2989.

Chubb, C. & Sperling, G. (1989b). Second-order motion perception: Space-time separable mechanisms. *Proceedings: Workshop on Visual Motion (March 20-22, 1989, Irvine, CA)* pp. 126-138. Washington, D.C.:IEEE Computer Society Press.

Chubb, C. & Sperling, G. (1991). Texture quilts: Basic tools for studying motion-from-texture. *Journal of Mathematical Psychology,* 35, 411-442.

Cropper, S. J. (1994). Velocity discrimination in chromatic gratings and beats *Vision Research,* 34, 41-48.

Cropper, S.J. & Badcock, D.R. (1995). Appended at end of reference list, page 236.

Cropper, S.J. & Derrington, A.M. (1993). Motion of chromatic stimuli: first-order or second-order. *Vision Research,* 34, 49-58.

Cropper, S.J. & Derrington, A.M. (1996). Detection and motion detection in chromatic and luminance beats. *Journal of the Optical Society of America A,* 13, 401-407.

Cropper, S.J. & Hammett, S.T. (1996). No motion aftereffect for a second-order stimulus. *Investigative Opthamology & Visual Science (Suppl.),* 37, 900.

Davis, J. & Shah, M. (1994). Visual gesture recognition. *IEE Proceedings - Visual Image and Signal Processing*, 141, 101-106.

Derrington, A.M. (1987). Distortion products in geniculate X-cells: a physiological basis for masking by spatially modulated gratings? *Vision Research*, 27, 1377-1386.

Derrington, A.M. (1990). Mechanisms for coding luminance patterns: are they really linear? In Blakemore, C. (Ed). *Vision: coding and efficiency*. (pp. 175-184). Cambridge: Cambridge University Press.

Derrington, A.M. & Badcock, D.R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research*, 25, 1869-1878.

Derrington, A.M. & Badcock, D.R. (1986). Detection of spatial beats: non-linearity or contrast increment detection? *Vision Research*, 26, 1767-1776.

Derrington, A.M., Badcock, D.R. & Henning, G.B. (1993). Discriminating the direction of 2nd-order motion at short stimulus durations. *Vision Research*, 33, 1785-1794.

Derrington, A.M. & Henning, G.B. (1993). Detecting and discriminating the direction of luminance and colour gratings. *Vision Research*, 33, 799-811.

Derrington, A.M. & Henning, G.B. (1994). Implications of motion detection for early non-linearities. In Bock, G.R. & Goode, J.A. (Eds), *Higher order processing in the visual system: Ciba Foundation Symposium, 184.* (pp.211-226). Chichester: Wiley.

De Bruyn, B. & Orban, G. A. R. (1988). Human velocity and direction discrimination measured with random dot patterns *Vision Research,* 28, 1323-1335.

DeValois, R.L., Albrecht, D.G. & Thorell, L.G. (1982). Spatial frequency selectivity of cells in the macaque visual cortex. *Vision Research*, 22, 545-559.

DeValois, R.L., Yund, E.W. & Hepler, H. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22, 531-544.

Dobkins, K.R. & Albright, T.D. (1995). Behavioural and neural effects of chromatic isoluminance in the primate visual-motion system. *Visual Neuroscience*, 12, 321-332.

Drasdo, N. (1991). Neural substrates and threshold gradients of peripheral vision. In Kulikowski, J.J., Walsh, V. & Murray, I.J. (eds). *Limits of vision* (pp. 251-265). Basingstoke: Macmillan.

Edwards, M. & Badcock, M. (1995). Global motion perception: no interaction between the first- and second-order motion pathways. *Vision Research*, 35, 2589-2602.

Fahle, M. & Poggio, T. (1981). Visual hyperacuity: spatiotemporal interpolation in human vision. *Proceedings of the Royal Society of London B,* 213, 451-477.

Fennema, C.L. & Thompson, W.B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9, 301-315.

Findlay, J.M. (1978). Estimates on probability functions: A more virulent PEST. *Perception & Psychophysics*, 23, 181-185.

Finney, D.J. (1971). *Probit Analysis*. Cambridge: Cambridge University Press.

Fleet, D. J. & Langley, K. (1994). Computational analysis of non-Fourier motion. *Vision Research,* 34, 3057-3079.

Frisch, K. von. (1967). The dance language and orientation of bees. *Oxford: Oxford University Press.*

Gee, A. & Cipolla, R. (1994). Determining the gaze of faces in images. *Image and Vision Computing*, 12, 639-647.

Gegenfurtner, K.R. & Hawken, M.J. (1996a). Interactions of motion and colour in the visual pathways. *Trends in Neurosciences*, 19, 394-401.

Gegenfurtner, K.R. & Hawken, M.J. (1996b). Perceived velocity of luminance, chromatic and non-Fourier stimuli - influence of contrast and temporal frequency. *Vision Research*, 36, 1281-1290.

Gegenfurtner, K.R., Kiper, D.C., Beusmans, J.M.H., Carandini, M., Zaidi, Q. & Movshon, J.A. (1994). Chromatic properties of of neurons in macaque MT. *Visual Neuroscience*, 11, 455-466.

Georgeson, M.A. (1992). Human vision combines oriented filters to compute edges. *Proceedings of the Royal Society of London B*, 249, 235-245.

Georgeson, M.A. & Harris, M.G. (1990). The temporal range of motion sensing and motion perception. *Vision Research*, 30, 615-619.

Georgeson, M.A. & Shackleton, T.M. (1989). Monocular motion sensing, binocular motion perception. *Vision Research*, 29, 1511-1523.

Georgeson, M.A. & Shackleton, T.M. (1992). No evidence for dichoptic motion sensing: a reply to Carney and Shadlen. *Vision Research*, 32, 193-198.

Ginsburg, A.P., Cannon, M.W. & Nelson, M.A. (1980). Suprathreshold processing of complex visual stimuli: Evidence for linearity in contrast perception. *Science*, 208, 619-621.

Goddard, N.H. (1989). The interpretation of visual motion: recognizing moving light displays. *Proceedings: Workshop on Visual Motion (March 20-22, 1989, Irvine, CA)* pp. 212-220. Washington, D.C.:IEEE Computer Society Press.

Gorea, A. (1995). Spatiotemporal characterization of a Fourier and non-Fourier motion system. *Vision Research*, 35, 907-914.

Green, M. (1981). Psychophysical relationships among mechanisms sensitive to pattern, motion and flicker. *Vision Research*, 21, 971-983.

Greenlee, M.W. & Smith, A.T. (1997). Detection and discrimination of first- and second-order motion in patients with unilateral brain damage. *Journal of Neuroscience*, 17, 804-818.

Greenlee, M.W. & Smith, A.T., Singh, K.D., Kraemer, F.M. & Hennig, J (1997). Functional echo-planar imaging of visual cortex during second-order motion perception. *Investigative Opthamology & Visual Science (Suppl.)*, 38, 238.

Gregory, R.L. (1966). Eye and Brain. *McGraw-Hill, New York.*

Grzywacz, N.M. & Yuille, A.L. (1990). A model for the estimate of local image velocity by cells in the visual cortex. *Proceedings of the Royal Society of London B,* 239, 129-161.

Hamerly, J.R., Quick, R.F., Jr. & Reichert, T.A. (1977). A study of grating contrast judgement. *Vision Research*, 17, 201-207.

Hammett, S.T., Ledgeway, T. & Smith, A.T. (1993). Transparent motion from feature-based and luminance-based processes. *Vision Research*, 33, 1119-1122.

Hammett, S.T. & Smith, A.T. (1994). Temporal beats in the human visual system. *Vision Research*, 34, 2833-2840.

Harris, M.G. (1980). Velocity specificity of the flicker to pattern sensitivity ration in human vision. *Vision Research*, 20, 687-691.

Harris, L.R. & Smith, A.T. (1992). Motion defined exclusively by 2nd-order characteristics does not evoke optokinetic nystagmus. *Visual Neuroscience*, 9, 565-570.

Heeger, D.J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A,* 4, 1455-1471.

Henning, G.B., Hertz, B.G. & Broadbent, D.E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research,* 15, 887-897.

Holliday, I.E. & Anderson, S.J. (1994). Different processes underlie the detection of second-order motion at low and high temporal frequencies. *Proceedings of the Royal Society of London B,* 257, 165-173.

Horn, B.K.P & Schunck, B.G. (1980). Determining optical flow. *Massachusetts Institute of Technology A.I. Memo No. 572.*

Hotson, J., Braun, D., Herzberg, W. & Boman, D. (1994). Transcranial magnetic stimuluation of extrastriate cortex degrades human motion direction discrimination. *Vision Research,* 34, 2115-2123.

Jamar, J.H.T. & Koenderink, J.J. (1985). Contrast detection and detection of contrast modulation for noise gratings, *Vision Research,* 25, 511-521.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics,* 14, 201-211.

Johnston, A. (1986). A spatial property of retinocortical mapping. *Spatial Vision,* 1, 319-331.

Johnston, A. (1987). Spatial scaling of central and peripheral contrast-sensitivity functions. *Journal of the Optical Society of America, A,* 4, 1583-1593.

Johnston, A. & Benton, C.P. (1996). Velocity discrimination thresholds for first- and second-order bars and edges. *Investigative Opthamology & Visual Science (Suppl.),* 37, 4167.

Johnston, A. & Clifford, C. W. G. (1995a). Perceived motion of contrast modulated gratings: Predictions of the Multi-channel Gradient Model and the role of full-wave rectification. *Vision Research,* 35, 1771-1783.

Johnston, A. & Clifford, C. W. G. (1995b). A unified account of three apparent motion illusions. *Vision Research,* 35, 1109-1123.

Johnston, A., McOwan, P. W. (1996). Velocity computation from measures of spatiotemporal gradients at multiple orientations. *Perception (Suppl.)*, 25, 35.

Johnston, A., McOwan, P. W. & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of the Royal Society of London B,* 250, 297-306.

Kukkonen, H., Rovamo, J., Tiipana, K. & Näsänen, R. (1993). Michelson contrast, RMS contrast and energy of various spatial stimuli at threshold. *Vision Research*, 33, 1431-1436.

Kukkonen, H., Rovamo, J. & Näsänen, R. (1995). Masking potency and whiteness of noise at various check sizes. *Investigative Opthalmology & Visual Science*, 36, 513-518.

Kulikowski, J.J. (1976). Effective contrast constancy and linearity of contrast sensation. *Vision Research*, 16, 1419-1431.

Langley, K. Fleet, D.J. & Hibbard, P.B. (1996). Linear filtering precedes nonlinear processing in early vision. *Current Biology*, 6, 891-896.

Ledgeway, T. (1994a). The detection of second-order motion in the human visual system. *PhD thesis, University of Cardiff.*

Ledgeway, T. (1994b). Adaptation to 2nd-order motion results in a motion aftereffect for directionally-ambiguous test stimuli. *Vision Research*, 34, 2879-2889.

Ledgeway, T. & Smith, A.T. (1994a). Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. *Vision Research*, 34, 2727-2740.

Ledgeway, T. & Smith, A.T. (1994b). The duration of the motion aftereffect following adaptation to first-order and second-order motion. *Perception*, 23, 1211-1219.

Ledgeway, T. & Smith, A.T. (1995). The perceived speed of second-order motion and its dependence on stimulus contrast. *Vision Research*, 35, 1421-1434.

Ledgeway, T. & Smith, A.T. (1996). Changes in perceived speed following adaptation to first-order and second-order motion. *Vision Research*, 37, 215-224.

Lee, D.N. & Lishman, J.R. (1975). Visual proprioceptive control of stance. *Journal of Human Movement Studies*. 1, 87-95.

Legge, G.E., Kersten, D. & Burgess, A.E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America A*, 4, 391-404.

Lu, Z-L. & Sperling, G. (1995a). Drastically different saturation for luminance motion versus texture-contrast motion. *Investigative Opthamology & Visual Science (Suppl.)*, 36, 395.

Lu, Z-L. & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Research*, 35, 2697-2722.

Lu, Z-L. & Sperling, G. (1995c). Attention-generated apparent motion. *Nature*, 377, 237-239.

Lu, Z-L. & Sperling, G. (1996). Second-order illusions: Mach bands, Chevreul, and Craik-O'Brien-Cornsweet. *Vision Research*, 36, 559-572.

Lucas, B.D. & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. *Proceedings of the 7th Joint International Conference on Artificial Intelligence.* Vancouver, B.C. 674-679.

MacLeod, D.I.A. & He, S. (1993). Visible flicker from invisible patterns. *Nature*, 361, 256-258.

MacLeod, D.I.A., Williams, D.R. & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research*, 32, 347-363.

Maffei, L. & Fiorentini, A. (1973). The visual cortex a a spatial frequency analyser. *Vision Research*, 13, 1255-1267.

Mareschal, I. & Baker, C.L., Jr. (1996). Slower prefered velocities to non-Fourier stimuli in visual cortex. *Investigative Opthamology & Visual Science (Suppl.)*, 37, 483.

Mareschal, I. & Baker, C.L., Jr. (1997). Bandpass temporal and spatial frequency responses to envelopes of second order stimuli in area 18 neurons. *Investigative Opthamology & Visual Science (Suppl.)*, 38, 624.

Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London, B*, 207, 187-217.

Marr, D. & Ullman, S. (1981). Directional selectivity and its use in early vision. *Proceedings of the Royal Society of London, B*, 211, 151-180.

Mather, G. & Murdoch, L. (1996). Second-order four-stroke apparent motion. *Investigative Opthamology & Visual Science (Suppl.)*, 37, 900.

Mather, G. & West, S. (1993). Evidence for second-order motion detectors. *Vision Research*, 33, 1109-1112.

McKee, S.P. (1981). A local mechanism for differential velocity detection. *Vision Research*, 21, 491-500.

McKee, S.P., Silverman, G. & Nakayama, K. (1986). Precise velocity discrimination despite random variation in temporal frequency *Vision Research*, 26, 609-620.

McKee, S.P. & Watamaniuk, S.N.J. (1994). The psychophysics of motion perception. In A. T. Smith & R. J. Snowden (Ed.), *Visual Detection of Motion* (pp. 85-114). London: Academic Press.

McOwan, P.W. & Johnston, A.J. (1995). The algorithms of natural vision: The Multi-channel Gradient Model. *Genetic Algorithms in Engineering Systems: Innovations and Applications. (September 12-14, 1995, Sheffield, UK)* pp. 319-324. London: Institution of Electrical Engineers.

McOwan, P.W. & Johnston, A.J. (1996b). Motion transparency arises from perceptual grouping: evidence from luminance and contrast modulation motion displays. *Current Biology,* 6, 1343-1346.

McOwan, P.W. & Johnston, A.J. (1996b). A second-order pattern reveals separate strategies for encoding orientation in two-dimensional space and space-time. *Vision Research,* 36, 425-430.

Merigan, W.H., Byrne, C.E. & Maunsell, J.H.R. (1991). Does primate motion perception depend on the magnocellular pathway? *Journal of Neuroscience*, 11, 3422-3429.

Merigan, W.H. & Maunsell, J.H.R. (1993). How parallel are the primate visual pathways. *Annual Review of Neuroscience*, 16, 369-402.

Moulden, B., Kingdom, F, & Gatley, L.F. (1990). The standard deviation of luminance as a metric for contrast in random-dor images. *Perception*, 19, 79-101.

Müller, R. & Greenlee, M.W. (1994). Effect of contrast and adaptation on the perception of the direction and speed of drifting gratings, *Vision Research*, 34, 2071-2092.

Nachmias, J. & Rogowitz, B.E. (1983). Masking by spatially-modulated gratings. *Vision Research*, 23, 1621-1629.

Newsome, W.T. & Paré, E.B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, 8, 2201-2211.

Newsome, W.T., Wurtz, R.H., Dürsteler, M.R. & Mikami, A. (1985). Deficits in motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *Journal of Neuroscience*, 5, 825-840.

Nishida, S. (1993). Spatiotemporal properties of motion perception for random-check contrast modulations. *Vision Research*, 33, 633-645.

Nishida, S., Ledgeway, T. & Edwards, M. (1997). Dual multiple-scale processing for motion in the human visual system. *Vision Research*, 37, 2685-2698.

Orban, G. A., De Wolf, J. & Maes, H. (1984). Factors influencing velocity coding in the human visual system *Vision Research*, 24, 33-39.

Orban, G. A., van Calenbergh, F., De Bruyn, B. & Maes, H. (1985). Velocity discrimination in central and peripheral visual field *Journal of the Optical Society of America A*, 2, 1836-1847.

Panish, S. C. (1988). Velocity discrimination at constant multiples of threshold contrast *Vision Research*, 28, 193-201.

Pantle, A. (1992). Immobility of some second-order stimuli in human peripheral vision. *Journal of the Optical Society of America A*, 9, 863-867

Pantle, A. & Turano, K. (1992). Visual resolution of motion ambiguity with periodic luminance- and contrast-domain stimuli. *Vision Research*, 32, 2093-2106.

Quick, R.F., Jr., Hamerly, J.R. & Reichert, T.A. (1976). The absence of a measure of "critical band" at low suprathreshold contrasts. *Vision Research*, 16, 351-355.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. *In Rosenblith, W.A. (Ed). Sensory Communication.* New York: Wiley.

Rizzo, M., Nawrot, M. & Zihl, J. (1995), Motion and shape pereption in cerebral akinotopsia. *Brain*, 118, 1105-1127.

Robson, J.G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America*, 56, 1141-1142.

Rodieck, R.W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, 5, 583-601.

Rovamo, J. & Kukkonen, H. (1996). The effect of noise check size and shape on grating detectability. *Vision Research,* 36, 271-279.

Rovamo, J. & Virsu. V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, 37, 495-510.

Sato, T. (1989). Reversed apparent motion with random dot patterns. *Vision Research*, 29, 1749-1758.

Scott-Samuel, N.E. (1997). Mechanisms of motion perception in human vision, 1st-order, 2nd-order and feature matching: psychophysical investigations. *PhD thesis, University of Birmingham.*

Scott-Samuel, N.E. & Georgeson, M.A. (1995). Does early nonlinearity account for second-order motion? *Perception (Suppl.)*, 24, 104.

Solomon, J. A. & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research*, 34, 2239-2258.

Soodak, R.E. (1986). Two-dimensional modeling of visual receptive fields using Gaussian subunits. *Proceedings of the National Academy of Sciences, U.S.A.*, 83, 9259-9263..

Smith, A.T. (1985). Velocity coding: evidence from perceived velocity shifts. *Vision Research*, 25, 1969-1976.

Smith, A. T. (1987). Velocity perception and discrimination: relation to temporal mechanisms *Vision Research*, 27, 1491-1500.

Smith, A.T. (1994a). The detection of second-order motion. In Smith, A.T. & Snowden, R.J. (Eds), *Visual Detection of Motion.* London: Academic Press.

Smith, A.T. (1994b). Correspondence-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society of America A,* 11, 1940-1948.

Smith, A. T. & Edgar, G. K. (1991). The separability of temporal frequency and velocity *Vision Research*, 31, 321-326.

Smith, A.T., Hess, R.F. & Baker, J.C.L. (1994). Direction identification thresholds for second- order motion in central and peripheral vision. *Journal of the Optical Society of America A,* 11, 506-514.

Smith, A.T. & Ledgeway, T. (1997a). Separate detection of moving luminance and contrast modulations: Fact or Artifact? *Vision Research*, 37, 45-62.

Smith, A.T. & Ledgeway, T. (1997b). Sensitivity to second-order motion as a function of drift temporal frequency and viewing eccentricity. *Investigative Opthamology & Visual Science (Suppl.)*, 38, 82.

Smith, E.L., Chino, Y.M., Ridder, W.H., Kitagawa, K. & Langston, A. (1990). Orientation bias of neurons in the lateral geniculate-nucleus of macaque monkeys. *Visual Neuroscience*, 5, 525-545.

Snowden, R. J. & Braddick, O.J. (1991). The temporal integration and resolution of velocity signals *Vision Research*, 31, 907-914.

Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision,* 4, 183-207.

Sperling, G., Chubb, C., Solomon, J.A. & Lu, Z-L. (1994). Full-wave and half-wave processes in second-order motion and texture. In Bock, G.R. & Goode, J.A. (Eds), *Higher order processing in the visual system: Ciba Foundation Symposium, 184.* (pp.287-303). Chichester: Wiley.

Sperling, G. & Lu, Z-L. (1996). Second-order reversed-phi reveals two mechanisms: second-order motion energy and third-order feature-salience. *Investigative Opthamology & Visual Science (Suppl.)*, 37, 900.

Stone, L.S. & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, 32, 1535-1549.

Sutter, A., Sperling, G. & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research*, 35, 915-924.

Taylor, M.M. & Creelman, C.D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, 41, 782-787.

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, 22, 377-380.

Thompson, P. (1983). Discrimination of moving gratings at and above detection threshold. *Vision Research*, 23, 1533-1538.

Thompson, P., Stone, L.S. & Swash, S. (1996). Speed estimates from grating patches are not contrast-normalized. *Vision Research*, 36, 667-674.

Tiippana, K., Näsänen, R. & Rovamo, J. (1994). Contrast matching of two-dimensional compound gratings. *Vision Research*, 34, 1157-1163.

Tomita, T. (1968). Electrical responses of single photoreceptors. *Proceedings of the IEEE*, 56, 1015-1023.

Tootell, R.B.H, Reppas, J.B., Kwong, K.K, Malach, R., Born, R.T., Brady, T.J., Rosen, B.R. & Belliveau, J.W. (1995). Function analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, 15, 3215-3230.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 35, 2503-2522.

Turano, K. (1991). Evidence for a common motion mechanism of luminance-modulated and contrast-modulated patterns: selective adaptation. *Perception*, 20, 455-466.

Turano, K. & Pantle, A. (1989). On the mechanism that encodes the movement of contrast variations. I: velocity discrimination. *Vision Research*, 29, 207-221.

Vaina, L.M. & Cowey, A. (1996). Impairment of the perception of second-order motion but not first order motion in a patient with unilateral brain damage. *Proceedings of the Royal Society of London B*, 263, 1225-1232.

van Essen, D.C., Maunsell, J.H.R. & Bixby, J.L. (1981). The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *Journal of Comparative Neurology*, 199, 293-326.

van Santen, J.P.H. & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, 1, 451-.473.

van Santen, J.P.H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2, 300-321.

Verstraten, F.A.J. (1996). On the ancient history of the direction of the motion aftereffect. *Perception*, 25, 1177-1187.

Victor, J.D. & Conte, M.M. (1992). Evoked-potential and psychophysical analysis of Fourier and non-Fourier motion mechanisms. *Visual Neuroscience*, 9, 105-123.

Watamaniuk, S. N. J., Grzwacz, N. M. & Yuille, A. L. (1993). Dependence of speed and direction perception on cinematogram dot density *Vision Research*, 33, 849-857.

Watson, A.B. & Ahumada, A.J., Jr. (1983). A look at motion in the frequency domain. *NASA Technical Memorandum 84352*.

Watson, A.B. & Ahumada, A.J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2, 322-342.

Watson, A.B. & Robson, J.G. (1981). Discrimination at threshold: labelled detectors in human vision, *Vision Research*, 21, 1115-1122.

Watson, J.D.G., Myers, R., Frackowiak, R.S.J, Hajnal, J.V., Woods, R.P., Mazziotta, J.C., Shipp, S. & Zeki, S. (1993). Area V5 of the human brain: Evidence form a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, 3, 79-94.

Watt, R. J. & Andrews, D. P. (1981). APE: Adaptive Probit Estimation of psychometric functions *Current Psychological Reviews,* 1, 205-214.

Watt, R.J. & Morgan, M.J. (1985). A theory of the primitive spatial code in human vision. *Vision Research,* 25, 1661-1647.

Wilson, H.R. (1994). Models of two-dimensional motion perception. In Smith, A.T. & Snowden, R.J. (Eds), *Visual Detection of Motion.* London: Academic Press.

Wilson, H.R., Ferrara, V.P. & Yo, C. (1992). Psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience,* 9, 79-97.

Wohlgemuth, A. (1911). On the aftereffect of seen movement. *British Journal of Psychology, Monographs 1.*

Yo, C. & Wilson, H.R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research,* 32, 135-147.

Zeki, S. (1974). Functional organization of a visual area in the posterior bank of the temporal sulcus of the rhesus monkey. *Journal of Physiology, London,* 236, 549-573.

Zeki, S. (1993). A vision of the brain. *Oxford: Blackewell Scientific Publications.*

Zeki, S., Watson, J.D.G, Lueck, C.J., Friston, K.J., Kennard, C. & Frackowiak, R.S.J. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience,* 11, 641-649.

Zhou, Y-X, & Baker, C.L., Jr. (1993). A processing stream in the mammalian visual cortex neurons for non-Fourier responses. *Science,* 261, 98-101.

Zhou, Y-X, & Baker, C.L., Jr. (1994). Envelope-responsive neurons in areas 17 and 18 of cat. *Journal of Neurophysiology,* 72, 2134-2150.

Zihl, J., von Cramon, D. & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain*, 106, 313-340.

Cropper, S.J. & Badcock, D.R. (1995). Perceived direction of motion: it takes all orientations. *Perception (Suppl.)*, 24, 106.